Ph.D. DISSERTATION

Improvement of FrodoKEM System by BCH Codes and Minimax Approximation of Sign Function by Composite Polynomial for Homomorphic Comparison

BCH 부호를 이용한 FrodoKEM의 성능 개선 및 동형 비교를 위한 합성함수에 의한 부호 함수의 미니맥스 근사

BY

EUNSANG LEE

AUGUST 2020

DEPARTMENT OF ELECTRICAL ENGINEERING AND
COMPUTER SCIENCE
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

Ph.D. DISSERTATION

Improvement of FrodoKEM System by BCH Codes and Minimax Approximation of Sign Function by Composite Polynomial for Homomorphic Comparison

BCH 부호를 이용한 FrodoKEM의 성능 개선 및 동형 비교를 위한 합성함수에 의한 부호 함수의 미니맥스 근사

BY

EUNSANG LEE

AUGUST 2020

DEPARTMENT OF ELECTRICAL ENGINEERING AND
COMPUTER SCIENCE
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

Improvement of FrodoKEM System by BCH Codes and Minimax Approximation of Sign Function by Composite Polynomial for Homomorphic Comparison

BCH 부호를 이용한 FrodoKEM의 성능 개선 및 동형 비교를 위한 합성함수에 의한 부호 함수의 미니맥스 근사

지도교수 노 종 선

이 논문을 공학박사 학위논문으로 제출함

2020년 8월

서울대학교 대학원

전기·정보공학부

이 은 상

이은상의 공학박사 학위 논문을 인준함

2020년 8월

위 원 장: _____
부위원장: _____
위 원: _____
위 원: _____
위 원: _____

Abstract

In this dissertation, two main contributions are given as;

(i) Performance improvement of FrodoKEM using Gray and error-correcting codes (ECCs).

(ii) Optimal minimax polynomial approximation of sign function by composite polynomial for homomorphic comparison.

First, modification of FrodoKEM using Gray codes and ECCs is studied. Lattice-based scheme is one of the most promising schemes for post-quantum cryptography (PQC). Among many lattice-based cryptosystems, FrodoKEM is a well-known key-encapsulation mechanism (KEM) based on (plain) learning with errors problems and is advantageous in that the hardness is based on the problem of unstructured lattices. Many lattice-based cryptosystems adopt ECCs to improve their performance, such as LAC, Three Bears, and Round5 which were presented in the NIST PQC Standardization Round 2 conference. However, for lattice-based cryptosystems that do not use ring structures such as FrodoKEM, it is difficult to use ECCs because the number of transmitted symbols is small. In this dissertation, I propose a method to apply Gray and ECCs to FrodoKEM by encoding the bits converted from the encrypted symbols. It is shown that the proposed method improves the security level and/or the bandwidth of FrodoKEM, and 192 message bits, 50% more than the original 128 bits, can be transmitted using one of the modified Frodo-640's.

Second, an optimal minimax polynomial approximation of sign function by a composite polynomial is studied. The comparison function of the two numbers is one of the most commonly used operations in many applications including deep learning and data processing systems. Several studies have been conducted to efficiently evaluate the comparison function in homomorphic encryption schemes which only allow ad-

dition and multiplication for the ciphertext. Recently, new comparison methods that approximate sign function using composite polynomial in the homomorphic encryption, called homomorphic comparison operation, were proposed and it was proved that the methods have optimal asymptotic complexity. In this dissertation, I propose new optimal algorithms that approximate the sign function in the homomorphic encryption by using composite polynomials of the minimax approximate polynomials, which are constructed by the modified Remez algorithm. It is proved that the number of required non-scalar multiplications and depth consumption for the proposed algorithms are less than those for any methods that use a composite polynomial of component polynomials with odd degree terms approximating the sign function, respectively. In addition, an optimal polynomial-time algorithm for the proposed homomorphic comparison operation is proposed by using dynamic programming. As a result of numerical analysis, for the case that I want to minimize the number of non-scalar multiplications, the proposed algorithm reduces the required number of non-scalar multiplications and depth consumption by about 33% and 35%, respectively, compared to those for the previous work. In addition, for the case that I want to minimize the depth consumption, the proposed algorithm reduces the required number of non-scalar multiplications and depth consumption by about 10% and 47%, respectively, compared to those for the previous work.

keywords: Error-correcting codes, FrodoKEM, fully homomorphic encryption (FHE), Gray code, homomorphic comparison operation, lattice-based cryptography, minimax approximation, post-quantum cryptography, Remez algorithm, sign function

student number: 2014-22573

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Existing public-key cryptosystems such as RSA and elliptic curve cryptography can be broken by future quantum computers because of the rapid development of quantum computers. In addition, quantum computers can perform an exhaustive search in the square root of time complexity of classical computers, and the key length of symmetric ciphers such as advanced encryption standard (AES) should be doubled for the same security level. Therefore, it is very important to develop secure post-quantum cryptography (PQC) algorithms resistant to quantum computing. In the first-round evaluation of PQC schemes submitted to NIST PQC Standardization (i.e., NIST PQC Round 1), 26 algorithms have been selected for NIST PQC Round 2 and are under evaluation. Among the algorithms selected for NIST PQC Round 2 are the lattice-based schemes, code-based schemes, multi-variate schemes, and so on. Notably, 12 of the 26 algorithms are lattice-based ones [1]–[3]. Thus, lattice-based cryptography is the most promising field for PQC [4, 5].

Learning with errors (LWE) is a problem that was first introduced in [6] and it is proved that LWE is more difficult to solve than some well-known mathematical problems on lattices. Cryptosystems such as Diffie-Hellman and RSA algorithms are

based on the hardness of integer factorization problem or discrete logarithm problem. However, Shor proposed an efficient quantum algorithm on a quantum computer to solve these problems [7]. On the other hand, efficient quantum algorithms that solve LWE have not been found yet. Thus, there are many proposed PQC schemes based on LWE. FrodoKEM is a lattice-based algorithm based on the hardness of LWE, and FrodoKEM is one of the representative lattice-based PQC schemes selected for the NIST PQC Round 2 [1].

In addition, a lot of research has been done on applying error-correcting codes (ECCs) to lattice-based PQC schemes to improve their performance. NewHope, Round5, LAC, and Three Bears are lattice-based schemes selected for NIST PQC Round 2 that use ECCs. NewHope [2] uses a simple error correction technique, called additive threshold encoding (ATE). Round5 [8], which is a combined algorithm of Hila5 [9] and Round2 [10], uses an ECC, called XE5 which is resistant to side-channel attacks. LAC [11, 12] uses BCH codes of large code lengths. Three Bears [13] uses BCH codes such that constant-time implementation is possible, but the performance of BCH codes used in Three Bears is relatively worse compared to the performance of BCH codes used in LAC. A lattice-based key-encapsulation mechanism (KEM) scheme called KCL [14], which was submitted to NIST PQC Round 1 but was not selected for NIST PQC Round 2, uses a single-error correcting code, lattice code in $\tilde{D}_4$ [15], or lattice code in E_8 [14]. In [16], the performance of NewHope was improved by using ECCs. The ATE technique used in NewHope is replaced by BCH codes or concatenated coding schemes of low-density parity check (LDPC) and BCH codes to improve the security level. However, for lattice-based cryptosystems that do not use ring structures such as FrodoKEM, it is difficult to use ECCs because the number of transmitted symbols is small.

Homomorphic encryption (HE) is a cryptographic system that allows an untrusted worker to perform algebraic operations over the encrypted data without learning anything about the data [17]. Figure 1.1 shows the description of HE. Due to this feature,

HE has been extensively studied and has attracted significant attention in various applications such as deep learning, medical data processing, etc. Until Gentry's seminal work [18] in 2009, HE schemes were able to perform only a few specific operations for the encrypted data. A fully homomorphic encryption (FHE) scheme which allows arbitrary computations on the encrypted data was first developed in [18] and many FHE schemes have since been proposed to increase the efficiency of homomorphic computation [19]–[22]. Recently, Cheon-Kim-Kim-Song (CKKS) and the fast fully homomorphic encryption over the torus (TFHE) are known as two typical FHE schemes.

Since FHE only provides addition and multiplication, it is usually difficult to perform non-polynomial operations. Among non-polynomial operations, the comparison operation is one of the most commonly used operations in actual applications, along with addition and multiplication. Therefore, some studies on homomorphic comparison operation have been done [23]. Recently, a method of performing the homomorphic comparison operation with optimal asymptotic complexity using composite polynomial has been studied [24].

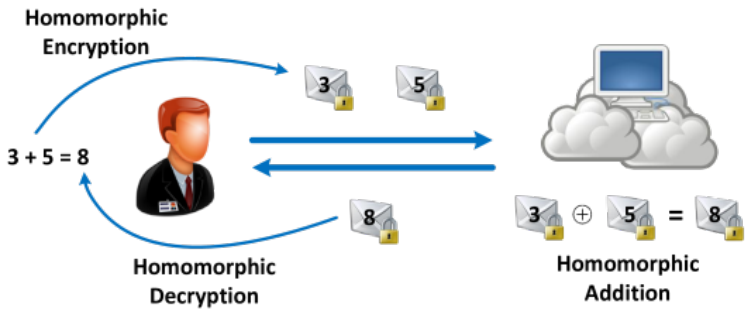

Figure 1.1: Description of homomorphic encryption.

1.2 Overview of Dissertation

This dissertation is organized as follows.

In Chapter 2, some preliminaries are presented. In Section 2.1, the NIST PQC standardization process is introduced and public-key encryption (PKE) and KEM schemes are described in Section 2.2. In Section 2.3, lattice-based cryptography is introduced, and BCH and Gray codes are described in Section 2.4. In Section 2.5, FHE is described and in Section 2.6, the concepts of approximation theory are introduced. Finally, algorithms for minimax approximation are described in Section 2.7.

In Chapter 3, modifying FrodoKEM with Gray codes and ECCs is proposed. In Section 3.1, FrodoPKE is described as a digital communication system and the method of modification of FrodoKEM with Gray codes and ECCs is proposed. In addition, ECCs and Gray codes to be used for FrodoPKE are designed and the method of calculating DFR is introduced. In Section 3.2, the results of the improvement of FrodoKEM by using Gray codes and ECCs are presented. First, the security level is improved by increasing the standard deviation of error. Second, the message size is increased. Finally, the bandwidth is reduced by decreasing the modulus.

In Chapter 4, applying a composition of minimax approximate polynomials for homomorphic comparison operation is proposed. Introduction is given in Section 4.1. In Section 4.2, a new method to approximate sign function using composite polynomial of minimax approximate polynomials is proposed. In addition, it is proved that the composite polynomials of minimax approximate polynomials obtained from the proposed method are optimal with respect to non-scalar multiplications or depth consumption among all the composite polynomial of polynomials with odd degree terms. Finally, achieving a polynomial-time algorithm for the homomorphic comparison by using dynamic programming is described. In Section 4.3, the numerical results of the improved homomorphic comparison operation by using the proposed algorithms are presented for both when the number of non-scalar multiplications is minimized and when the depth consumption is minimized. The required number of non-scalar multiplications and depths for the proposed algorithms are compared to those for the previous algorithm.

Finally, the concluding remarks are given in Chapter 5.

1.3 Notations

The following notations are used in this dissertation.

- For a finite set S, the uniform distribution on S is denoted by $U(S)$.
- For a probability distribution χ, drawing e value according to χ is denoted as $e \leftarrow \chi$.
- The ring of integers is denoted by $\mathbb{Z}$, and for a positive integer q, the quotient ring of integers modulo q is denoted by $\mathbb{Z}_q = \mathbb{Z}/q\mathbb{Z}$.
- For a real function $f(x)$, its infinity norm is denoted by $||f(x)||_\infty$.

Chapter 2

Preliminaries

In this chapter, some preliminaries are introduced. First, NIST PQC standardization is introduced. Second, the basic concepts of PKE and KEM schemes are given. Third, lattice-based cryptography is described. Fourth, the basic concepts of BCH and Gray codes are given. Fifth, the basic concept of FHE and comparison operation in FHE are introduced. Finally, the concepts of approximation theory and some algorithms that obtain the minimax approximate polynomials are given.

2.1 NIST Post-Quantum Cryptography Standardization

2.1.1 Background

Recently, a lot of research has been conducted on quantum computers. Quantum computers are machines that use quantum phenomena to solve mathematical problems that classical computers cannot solve. If large-scale quantum computers are developed, many commercial cryptographic systems will be broken. In particular, many public-key cryptographic systems such as RSA, elliptic curve cryptosystems, and digital signature algorithms, which are used for key establishment protocol and digital signatures, will be completely broken. With these concerns, many researchers have begun to study PQC schemes. The goal is to develop a secure cryptographic system against

quantum and classical computers. Developed PQC schemes will be replacements for commercial public-key cryptographic systems when large-scale quantum computers are developed.

There are several PQC schemes. Most of them are lattice-based cryptosystems, code-based cryptosystems, multi-variate cryptosystems, and hash signature system. However, more research is needed to gain more confidence in the security and to improve performance. Thus, NIST decided to start developing PQC standard for the following two reasons:

(i) There has been considerable progress in quantum computer research.

(ii) The transition to PQC seems to be very difficult. Significant efforts will be needed for developing, standardizing, and deploying new PQC cryptosystems. In addition, this transition should be done long before the development of large-scale quantum computers.

Several rounds are expected to take place for 3 to 5 years. The goal of the standardization process is to select many acceptable candidate algorithms. NIST expects that this evaluation process will be much more complex than the standardization processes of SHA-3 or AES. This is because public-key cryptosystems are more complicated than hash functions or block ciphers, and the current understanding of quantum computers is low.

2.1.2 Categories for Security Level

There is significant uncertainty in the PQC standardization process since it is difficult to predict the performance of quantum algorithms in the future. Thus, NIST defines a broad set of security categories from Category 1 to Category 5. Since, PKE and KEM schemes are only related to security categories 1, 3, and 5, only the security categories 1, 3, and 5 are introduced below. Figures 2.1 and 2.2 show the security categories and the required circuit sizes to break AES.

Table 2.1: NIST security categories.

	security description
I	At least as hard to break as AES128 (exhaustive key search)
III	At least as hard to break as AES192 (exhaustive key search)
V	At least as hard to break as AES256 (exhaustive key search)

Table 2.2: Required circuit sizes to break AES.

	security description
AES 128	2^{170} quantum gates or 2^{143} classical gates
AES 192	2^{233} quantum gates or 2^{207} classical gates
AES 256	2^{298} quantum gates or 2^{272} classical gates

2.1.3 List of Algorithms in NIST PQC Round 2

A total of 69 PQC algorithms were submitted to NIST PQC Round 1. The 26 algorithms among the 69 algorithms were selected for NIST PQC Round 2. Several kinds of PQC schemes such as lattice-based schemes, code-based schemes, multi-variate schemes, hash-based schemes, and an isogeny scheme were selected for NIST PQC Round 2. Among them, the 12 algorithms are lattice-based schemes. Thus, it can be seen that the most promising PQC is lattice-based cryptography. Table 2.3 shows the algorithms selected for NIST PQC Round 2.

Table 2.3: PQC algorithms accepted for NIST PQC Round 2.

	signatures	KEM / encryption
lattice-based	CRYSTALS-DILITHIUM	CRYSTALS-KYBER
	FALCON	FrodoKEM
	qTESLA	LAC
		NewHope
		NTRU
		NTRU Prime
		Round5
		SABER
		Three Bears
code-based		BIKE
		Classic McEliece
		HQC
		LEDAcrypt
		NTS-KEM
		ROLLO
		RQC
multi-variate	GeMss	
	LUOV	
	MQDSS	
	Rainbow	
hash-based	SPHINCS+	
isogeny		SIKE
finite automata	Picnic	

2.2 Public-Key Encryption and Key-Encapsulation Mechanism

PKE and KEM schemes are defined as follows. Figure 2.1 shows the description of a PKE scheme.

Definition 2.1. *A PKE scheme is a tuple of three polynomial-time algorithms that satisfy the followings:*

- $KeyGen(\lambda) \rightarrow (pk, sk)$*;* $KeyGen$ *takes security parameter* λ *as an input and outputs public key* pk *and secret key* sk*.*
- $Enc(\mu, pk) \rightarrow ct$*;* Enc *takes a public key* pk *and a message* μ *as inputs, and outputs a ciphertext* ct *of* μ*.*
- $Dec(ct, sk) \rightarrow \mu'$ *or* $\perp$*;* Dec *takes a ciphertext* ct *and a secret key* sk *as inputs, and output a message* μ'*. If the decryption procedure fails,* Dec *outputs a special symbol* $\perp$*.*

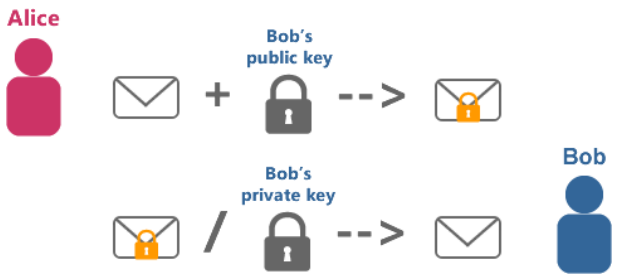

Figure 2.1: Description of a public-key encryption scheme.

Definition 2.2. *A KEM scheme is a tuple of three polynomial-time algorithms that satisfy the followings:*

- $KeyGen(\lambda) \rightarrow (pk, sk)$*;* $KeyGen$ *takes security parameter* λ *as an input and outputs public key* pk *and secret key* sk*.*

- $Encaps(pk) \rightarrow (ct, ss)$*; Encaps takes a public key* pk *and outputs an encapsulation* ct *and shared secret* ss*.*
- $Decaps(ct, sk) \rightarrow ss'$*; Decaps takes an encapsulation* ct *and a secret key* sk *as inputs, and outputs a shared secret* ss'*.*

Security notions of indistinguishability under chosen-plaintext attack (IND-CPA) and indistinguishability under chosen-ciphertext attack (IND-CCA) are now defined as follows.

Definition 2.3. *The chosen-plaintext attack (CPA) indistinguishability experiment* $PubK^{cpa}_{\mathcal{A},\Pi}(n)$ *is defined for PKE scheme* $\Pi = (KeyGen, Enc, Dec)$ *and adversary* $\mathcal{A}$ *as follows:*

(i) $KeyGen(\lambda)$ *obtains* (pk, sk)*.*

(ii) Adversary $\mathcal{A}$ *is given* pk *as well as oracle access to* $Enc(\cdot, pk)$*. The adversary outputs a pair of messages* μ_0 *and* μ_1 *with* $|\mu_0| = |\mu_1|$*.*

(iii) A random bit $b \in \{0, 1\}$ *is chosen, and the ciphertext* $ct \leftarrow Enc(\mu_b, pk)$ *is computed and given to* $\mathcal{A}$*.* ct *is called the challenge ciphertext.* $\mathcal{A}$ *continues to have access to* $Enc(\cdot, pk)$*.*

(iv) $\mathcal{A}$ *outputs a bit* b'*.*

(v) The output of the experiment is defined to be 1 if $b' = b$*, and 0 otherwise.*

Definition 2.4. *PKE scheme* $\Pi = (KeyGen, Enc, Dec)$ *is IND-CPA-secure if, for all probabilistic, polynomial-time adversaries* $\mathcal{A}$*, there exists a negligible function* $negl$ *such that*

$$Pr[PubK^{cpa}_{\mathcal{A},\Pi}(n) = 1] \leq \frac{1}{2} + negl(n).$$

Definition 2.5. *The chosen-ciphertext attack (CCA) indistinguishability experiment* $PubK^{cca}_{\mathcal{A},\Pi}(n)$ *is defined for PKE scheme* $\Pi = (KeyGen, Enc, Dec)$ *and adversary* $\mathcal{A}$ *as follows:*

(i) $KeyGen(\lambda)$ *obtains* (pk, sk).

(ii) Adversary $\mathcal{A}$ *is given* pk *and access to a decryption oracle* $Dec(\cdot, sk)$. *The adversary outputs a pair of messages* μ_0 *and* μ_1 *with* $|\mu_0| = |\mu_1|$.

(iii) A random bit $b \in \{0, 1\}$ *is chosen, and the ciphertext* $ct \leftarrow Enc(\mu_b, pk)$ *is computed and given to* $\mathcal{A}$.

(iv) $\mathcal{A}$ *continues to have access to the decryption oracle but can not request decryption of* ct *itself.*

(v) Finally, $\mathcal{A}$ *outputs a bit* b'.

(vi) The output of the experiment is defined to be 1 if $b' = b$, *and 0 otherwise.*

Definition 2.6. *PKE scheme* $\Pi = (KeyGen, Enc, Dec)$ *is IND-CCA-secure if, for all probabilistic, polynomial-time adversaries* $\mathcal{A}$, *there exists a negligible function* $negl$ *such that*

$$Pr[PubK^{cca}_{\mathcal{A},\Pi}(n) = 1] \leq \frac{1}{2} + negl(n).$$

In [25], it is shown that an IND-CCA-secure KEM can be constructed from an IND-CPA-secure PKE scheme and three hash functions by quantum Fujisaki Okamoto (QFO) transformation.

2.3 Lattice-Based Cryptography

2.3.1 Learning with Errors Problem

LWE problem was first introduced in [6]. χ is usually a discrete Gaussian of width αq for some $0 < \alpha < 1$. The definitions of LWE distribution, search-LWE problem, and decision-LWE problem are as follows.

Definition 2.7. *For a vector* $\mathbf{s} \in \mathbb{Z}_q^n$*, the LWE distribution* $A_{\mathbf{s},\chi}$ *over* $\mathbb{Z}_q^n \times \mathbb{Z}_q$ *is sampled by choosing* $\mathbf{a} \leftarrow U(\mathbb{Z}_q^n)$*, choosing* $e \leftarrow \chi$*, and outputting* $(\mathbf{a}, b = \langle \mathbf{s}, \mathbf{a} \rangle + e \bmod q)$.

Definition 2.8. *Search-LWE*$_{n,q,\chi,m}$ *problem is to find secret* $\mathbf{s}$ *for* m *given independent samples* $(\mathbf{a}_i, b_i) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$ *sampled from* $A_{\mathbf{s},\chi}$ *for a uniformly random* $\mathbf{s} \in \mathbb{Z}_q^n$.

Definition 2.9. *Decision-LWE*$_{n,q,\chi,m}$ *problem is to distinguish the case (with non-negligible advantage) for* m *independent samples* $(\mathbf{a}_i, b_i) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$ *of either: (1)* $A_{\mathbf{s},\chi}$ *for a uniformly random* $\mathbf{s} \in \mathbb{Z}_q^n$*, or (2) the uniform distribution.*

It is proved that Search-LWE and Decision-LWE are more difficult to solve than some well-known mathematical problems on lattices such as decisional approximate shortest vector problem (GapSVP) and shortest independent vectors problem (SIVP). In addition, efficient quantum algorithms that solve LWE have not been found yet. Thus, there are many PQC schemes based on LWE, which can be represented using a matrix as

$$\mathbf{b}^t = \mathbf{s}^t \mathbf{A} + \mathbf{e}^t \pmod{q}.$$

The columns of $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$ are the vectors $\mathbf{a}_i \in \mathbb{Z}_q^n$, the entries of the vector $\mathbf{b}$ are $b_i \in \mathbb{Z}_q$, and $\mathbf{e}$ is sampled from χ^m.

2.3.2 Overview of FrodoPKE Algorithm

FrodoKEM is one of the representative PQC cryptosystems that use LWE problem. This section reviews FrodoPKE, which is the underlying algorithm of FrodoKEM. FrodoKEM is the QFO transformation of FrodoPKE, and the decryption failure rate (DFR) performance of FrodoKEM is the same as that of FrodoPKE [1]. For simplicity, I apply ECCs to FrodoPKE instead of FrodoKEM. In FrodoKEM schemes submitted to PQC Round 1, there are two kinds of FrodoKEM schemes: Frodo-640 for security category 1 and Frodo-976 for security category 3. Another scheme, Frodo-1344, was also proposed for security category 5 in FrodoKEM schemes submitted to PQC Round 2. In this dissertation, I focus only on Frodo-640 and Frodo-976. These algorithms have the same form but use different parameters. The algorithms of FrodoPKE are described with the following parameters:

- χ; a probability distribution of approximated rounded Gaussian distribution with small support set defined on the set of integers, $\mathbb{Z}$
- q; a power-of-two integer modulus
- $\bar{m}, \bar{n}, n$; dimensions of matrices
- B; the number of bits per each symbol, where bits mean the codeword bits if ECC is used and the message bits, otherwise
- $len_{\mathbf{A}}$; the length of seeds for pseudorandom matrix generation for public key
- $len_{\mathbf{E}}$; the length of seeds for pseudorandom bit generation for error sampling

Algorithm 1: FrodoPKE.KeyGen [1]

Input: None

Output: Key pair $(pk, sk) \in (\{0,1\}^{len_{\mathbf{A}}} \times \mathbb{Z}_q^{n\times\bar{n}}) \times \mathbb{Z}_q^{n\times\bar{n}}$

$seed_{\mathbf{A}} \leftarrow U(\{0,1\}^{len_{\mathbf{A}}})$;

Generate pseudorandom matrix $\mathbf{A} \in \mathbb{Z}_q^{n\times n}$ with $seed_{\mathbf{A}}$ for public key;

$seed_{\mathbf{E}} \leftarrow U(\{0,1\}^{len_{\mathbf{E}}})$ for generation of error matrix $\mathbf{E}$;

Generate matrices $\mathbf{S}, \mathbf{E} \in \mathbb{Z}_q^{n\times\bar{n}}$ from $seed_{\mathbf{E}}$ according to χ distribution;

Compute $\mathbf{B} = \mathbf{AS} + \mathbf{E}$;

Return public key $pk \leftarrow (seed_{\mathbf{A}}, \mathbf{B})$ and secret key $sk \leftarrow \mathbf{S}$;

Algorithm 2: FrodoPKE.Enc [1]

Input: Message $\mu \in \{0,1\}^{\bar{m}\bar{n}B}$ and public key

$$pk = (seed_{\mathbf{A}}, \mathbf{B}) \in \{0,1\}^{len_{\mathbf{A}}} \times \mathbb{Z}_q^{n\times\bar{n}}$$

Output: Ciphertext $ct = (\mathbf{C}_1, \mathbf{C}_2) \in \mathbb{Z}_q^{\bar{m}\times n} \times \mathbb{Z}_q^{\bar{m}\times\bar{n}}$

Generate pseudorandom matrix $\mathbf{A} \in \mathbb{Z}_q^{n\times n}$ with $seed_{\mathbf{A}}$;

$seed_{\mathbf{E}} \leftarrow U(\{0,1\}^{len_{\mathbf{E}}})$;

Generate error matrices $\mathbf{S}', \mathbf{E}' \in \mathbb{Z}_q^{\bar{m}\times n}$ and $\mathbf{E}'' \in \mathbb{Z}_q^{\bar{m}\times\bar{n}}$ from $seed_{\mathbf{E}}$ according to χ distribution;

Compute $\mathbf{B}' = \mathbf{S}'\mathbf{A} + \mathbf{E}'$ and $\mathbf{V} = \mathbf{S}'\mathbf{B} + \mathbf{E}''$;

Return ciphertext $ct \leftarrow (\mathbf{C_1}, \mathbf{C_2}) = (\mathbf{B}', \mathbf{V} + Frodo.Encode(\mu))$

Algorithm 3: FrodoPKE.Dec [1]

Input: $\mathbf{C_1}, \mathbf{C_2}, \mathbf{S}$

Output: μ'

Compute $\mathbf{M} = \mathbf{C_2} - \mathbf{C_1}\mathbf{S} = \mathbf{V} + Frodo.Encode(\mu) - (\mathbf{S}'\mathbf{A} + \mathbf{E}')\mathbf{S} = Frodo.Encode(\mu) + \mathbf{S}'\mathbf{E} + \mathbf{E}'' - \mathbf{E}'\mathbf{S} = Frodo.Encode(\mu) + \mathbf{E}'''$;

Return $\mu' \leftarrow Frodo.Decode(\mathbf{M})$;

Bob generates a public key and a secret key through Algorithm 1 and sends the

public key to Alice. Alice generates ciphertexts $\mathbf{C}_1$ and $\mathbf{C}_2$ through Algorithm 2 with the received public key and the message μ and then sends them to Bob. Finally, Bob computes $\mathbf{M} = \mathbf{C_2} - \mathbf{C_1 S}$ and restores the message μ sent by Alice as in Algorithm 3.

The $Frodo.Encode$ function in Algorithm 2 is defined as follows. $Frodo.Encode$ takes a message $\mu \in \mathbb{Z}^{\bar{m}\bar{n}B}$ as input and outputs a matrix in $\mathbb{Z}_q^{\bar{m}\times\bar{n}}$. In Frodo-640 and Frodo-976, the values of B are 2 and 3, respectively, and B-bit messages are encoded into symbols in $\mathbb{Z}_q$ according to the following rules:

- $B = 2$ without Gray coding;

$$00 \to 0, 01 \to \frac{q}{4}, 10 \to \frac{2q}{4}, 11 \to \frac{3q}{4} \tag{2.1}$$

- $B = 3$ without Gray coding;

$$\begin{aligned} &000 \to 0, \quad 001 \to \frac{q}{8}, \quad 010 \to \frac{2q}{8}, \quad 011 \to \frac{3q}{8}, \\ &100 \to \frac{4q}{8}, \quad 101 \to \frac{5q}{8}, \quad 110 \to \frac{6q}{8}, \quad 111 \to \frac{7q}{8} \end{aligned} \tag{2.2}$$

The $Frodo.Decode$ function in Algorithm 3 is defined as follows. In Frodo-640 and Frodo-976, Bob rounds each component of the matrix $\mathbf{M}$ to the nearest multiples of $q/4$ or $q/8$, respectively. Then, Bob obtains the message μ' by applying the inverse of mapping in (2.1) or (2.2) to each rounded symbol in $\mathbb{Z}_q$.

Table 2.4: Initial error distributions χ in FrodoPKE [1].

	σ	0	±1	±2	±3	±4	±5	±6	±7	±8	±9	±10	±11
Frodo-640	2.75	$\frac{9756}{65536}$	$\frac{8857}{65536}$	$\frac{7280}{65536}$	$\frac{5249}{65536}$	$\frac{3321}{65536}$	$\frac{1844}{65536}$	$\frac{898}{65536}$	$\frac{384}{65536}$	$\frac{144}{65536}$	$\frac{47}{65536}$	$\frac{13}{65536}$	$\frac{3}{65536}$
Frodo-976	2.3	$\frac{11278}{65536}$	$\frac{10277}{65536}$	$\frac{7774}{65536}$	$\frac{4882}{65536}$	$\frac{2545}{65536}$	$\frac{1101}{65536}$	$\frac{396}{65536}$	$\frac{118}{65536}$	$\frac{29}{65536}$	$\frac{6}{65536}$	$\frac{1}{65536}$	

The initial error distributions χ of FrodoPKE are shown in Table 2.4 and are derived according to the following procedure. First, a Gaussian distribution is obtained with a given standard deviation σ. Next, a rounded Gaussian distribution is derived from it. Finally, the error distribution of the small support that approximates the rounded Gaussian distribution is obtained [1].

Let ψ be the product distribution of the two initial error distributions. Let χ' be the error distribution obtained by convolving ψ $2n$ times and then convolving the resulting distribution with χ. Each component of $\mathbf{E}'''$ in Algorithm 3 follows the distribution χ', and the standard deviation of χ' is approximately $\sigma' \approx \sigma\sqrt{2n\sigma^2 + 1}$.

2.3.3 Parameters of FrodoKEM

Table 2.5: Parameter sets of FrodoKEM [1].

	n	q	σ	B	$\bar{m} \times \bar{n}$	DFR	ct size (bytes)
Frodo-640	640	2^{15}	2.75	2	8×8	$2^{-148.8}$	9736
Frodo-976	976	2^{16}	2.3	3	8×8	$2^{-199.6}$	15768

Important FrodoKEM parameters are given in Table 2.5. Frodo-640 and Frodo-976 satisfy security categories 1 and 3 in the NIST PQC Standardization, respectively. How to compute the security level of FrodoKEM is described in [1]. In this dissertation, the FrodoKEM python source code supported by submitters of FrodoKEM is used to calculate the security level of various cases. Actually, the security level of the actual FrodoKEM is derived from a series of reductions, which is 5 or 6 bits smaller than the security level computed by the source code supported by submitters of FrodoKEM. Nevertheless, it is still meaningful to use this source code because my goal is not to obtain the accurate security level but to show improvement by using ECCs.

FrodoKEM has various parameters, $n, q, \sigma, B, \bar{m}$, and $\bar{n}$, which determine the bandwidth, computational complexity, the security level, and the DFR, respectively.

To satisfy security categories 1 and 3, it is recommended to set the security level higher than 143 and 207 bits, respectively. In addition, because there is an attack method by using decryption failure [26], the DFR should be low. Therefore, it is desirable that the DFRs are less than about 2^{-148} and 2^{-199} for security categories 1 and 3, respectively.

2.4 BCH and Gray Codes

BCH codes were developed in 1960 [27]. These codes can correct multiple errors and exhibit good error correction performance even for small code length. Relatively simple and feasible encoding and decoding techniques are also known, and hence, BCH codes have been widely used.

The code length of the BCH code is $n = q^m - 1$ for some prime q, and $q = 2$ holds for binary codes. In this dissertation, binary BCH codes are used. BCH codes are usually denoted by (l_n, l_k, l_t), where l_n is the length of codeword, l_k is the length of message, and l_t is the error-correction capability, i.e., the maximum number of correctable errors.

The Peterson-Gorenstein-Zierler decoding algorithm [28] has long been known for efficient decoding, and its complexity is $O(l_n l_t)$. BCH decoding algorithms do not usually have a constant computation time. However, using the method given in [29], constant-time decoding of the BCH code can be implemented to defend some side-channel attacks.

Gray code is a code designed to change only one binary bit for the adjacent symbols. The most commonly used binary Gray code is the reflected binary Gray code and it is also used in this dissertation. Gray code is used a lot in wireless communication systems. For higher-order modulations such as pulse amplitude modulation (PAM) or quadrature amplitude modulation (QAM), Gray code is used to lower bit error probability. Table 2.6 shows an example of a typical Gray code for 4 bits.

Table 2.6: An example of Gray code for 4 bits.

decimal	binary code	Gray code
0	0000	0000
1	0001	0001
2	0010	0011
3	0011	0010
4	0100	0110
5	0101	0111
6	0110	0101
7	0111	0100
8	1000	1100
9	1001	1101
10	1010	1111
11	1011	1110
12	1100	1010
13	1101	1011
14	1110	1001
15	1111	1000

2.5 Fully Homomorphic Encryption

2.5.1 Homomorphic Encryption

In the IoT era, a lot of devices communicate over the Internet. A third party will inevitably be asked to process the data because many devices cannot process data on their own. However, if the data to be processed is confidential and the third party is unreliable, the data should be sent encrypted, and the third party should perform operations on the encrypted data. HE allows operations over the encrypted data without decryption for this case.

Until Gentry's seminal work [18] in 2009, HE schemes were able to perform only a few specific operations on the encrypted data. FHE is a cryptosystem that can perform infinite number of algebraic operations on the encrypted data with bootstrapping. A FHE scheme was first developed in [18] and many FHE schemes have since been proposed to improve efficiency [20, 21, 22]. From now on, I will consider only the FHE rather than the HE.

FHE schemes are classified as bit-wise FHE and word-wise FHE. The basic operations of bit-wise FHE are logic gates, and the basic operations of word-wise FHE are algebraic operations such as addition and multiplication. In this dissertation, I focus only on word-wise FHE and thus the FHE is used instead of word-wise FHE. The definition of FHE is given as follows.

Definition 2.10. *A FHE scheme E is a set of five polynomial-time algorithms that satisfy the followings:*

- $\mathsf{KeyGen}(\lambda) \rightarrow (\mathrm{pk}, \mathrm{sk})$*;* KeyGen *takes security parameter λ as an input and outputs public key* pk *and secret key* sk*.*
- $\mathsf{Enc}(\mu, \mathrm{pk}) \rightarrow \mathrm{ct}$*;* Enc *takes a public key* pk *and a message μ as inputs, and outputs a ciphertext* ct *of μ.*
- $\mathsf{Dec}(\mathrm{ct}, \mathrm{sk}) \rightarrow \mu'$ *or* $\perp$*;* Dec *takes a ciphertext* ct *and a secret key* sk *as inputs, and outputs a message μ'. If the decryption procedure fails,* Dec *outputs a*

special symbol $\perp$.

- $\mathsf{Add}(\mathrm{ct}_1, \mathrm{ct}_2, \mathrm{evk})$*;* Add *takes ciphertexts* ct_1 *and* ct_2 *of* μ_1 *and* μ_2*, respectively, and an evaluation key* evk *as inputs, and outputs a ciphertext* $\mathrm{ct}_{\mathrm{add}}$ *of* $\mu_1 + \mu_2$.
- $\mathsf{Mult}(\mathrm{ct}_1, \mathrm{ct}_2, \mathrm{evk})$*;* Mult *takes ciphertexts* ct_1 *and* ct_2 *of* μ_1 *and* μ_2*, respectively, and an evaluation key* evk *as inputs, and outputs a ciphertext* $\mathrm{ct}_{\mathrm{mult}}$ *of* $\mu_1 \cdot \mu_2$.

In CKKS scheme, there are two kinds of multiplications: scalar multiplication and non-scalar multiplication. Non-scalar multiplications require much more computational complexity than scalar multiplications. Thus, in this dissertation, when the homomorphic comparison operation is considered, I focus on reducing the number of non-scalar multiplications rather than scalar multiplications, together with depth consumption.

2.5.2 Comparison Operation in Fully Homomorphic Encryption

FHEs support addition and multiplication operations on the encrypted data, but do not support any non-arithmetic operations such as comparison operation. Thus, the approximation of comparison operation should be performed by using addition and multiplication operations in FHE. The comparison function and sign function are denoted as

$$\mathrm{comp}(a,b) = \begin{cases} 1 & \text{if } a > b \\ 1/2 & \text{if } a = b \\ 0 & \text{if } a < b \end{cases}, \quad \mathrm{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}.$$

My goal is to perform approximation for $\mathrm{comp}(a,b)$, which is implemented only with additions and multiplications. Note that $\mathrm{comp}(a,b)$ and $\mathrm{sgn}(x)$ functions have the following relationships as

$$\mathrm{sgn}(x) = 2\mathrm{comp}(x, 0) - 1,\ \mathrm{comp}(a, b) = \frac{\mathrm{sgn}(a-b)+1}{2}.$$

Thus, the approximation of $\mathrm{comp}(a, b)$ is equivalent to that of $\mathrm{sgn}(x)$. Therefore, I only focus on the polynomial approximation for $\mathrm{sgn}(x)$.

Even though the efficiency of FHEs has been improved a lot since the first FHE was developed in 2009, it is known that the non-scalar multiplication operation still takes a lot of computational complexity. In addition, since bootstrapping requires a lot of computational complexity, minimizing the depth consumption for the homomorphic comparison operation is also important, which reduces the number of bootstrappings. Thus, it is necessary to approximate $\mathrm{sgn}(x)$ by polynomials while minimizing the number of non-scalar multiplications and depth consumption.

Definition 2.11 ([24]). *For $\alpha > 0$ and $0 < \epsilon < 1$, a polynomial p is said to be (α, ϵ)-close to $\mathrm{sgn}(x)$ over $[-1, 1]$ if p satisfies the following:*

$$||p(x) - \mathrm{sgn}(x)||_{\infty,[-1,-\epsilon]\cup[\epsilon,1]} \leq 2^{-\alpha},$$

where $||\cdot||_{\infty,D}$ denotes the infinity norm over the domain D.

$\mathrm{sgn}(x)$ is discontinuous at $x = 0$, and thus it is impossible to exactly approximate $\mathrm{sgn}(x)$ near $x = 0$. Definition 2.11 means that the approximation error is guaranteed below $2^{-\alpha}$ only for $\epsilon \leq |x| \leq 1$. Figure 2.2 shows an example of a function satisfying (α, ϵ)-close.

2.6 Approximation Theory

In this section, some concepts for approximation theory are introduced.

Definition 2.12. *Let D be a closed subset of $[a, b]$. Let f be a continuous function on D. A polynomial p is said to be the minimax approximate polynomial of degree at most*

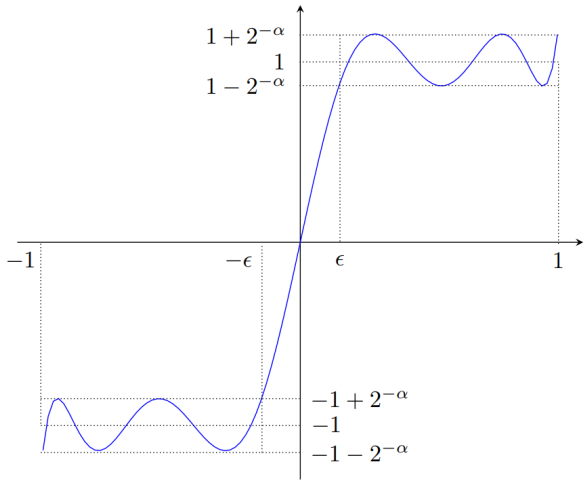

Figure 2.2: An example of an approximate polynomial satisfying (α, ϵ)-close for sign function.

n on D for f if p minimizes $\max_D ||p(x) - f(x)||_\infty$ among polynomials of degree at most n.

It is known that for any continuous function f on D, the minimax approximate polynomial of degree at most n on D for f uniquely exists [30]. I put $f(x) = \mathrm{sgn}(x)$ since the goal in this dissertation is to approximate $\mathrm{sgn}(x)$. I also only deal with cases where D is the union of two symmetric closed intervals, $[-b, -a] \cup [a, b]$.

Definition 2.13 (Haar's Condition and Generalized Polynomial [30])**.** *A set of functions $\{g_1, g_2, \cdots, g_n\}$ satisfies the Haar's condition if each g_i is continuous function and if the determinant*

$$D[x_1, \cdots, x_n] = \begin{vmatrix} g_1(x_1) & \cdots & g_n(x_1) \\ \vdots & \ddots & \vdots \\ g_1(x_n) & \cdots & g_n(x_n) \end{vmatrix}$$

is not zero for any n distinct points $x_1, \cdots, x_n$. A linear combination of $\{g_1, \cdots, g_n\}$ is referred to as a generalized polynomial.

The following theorem and lemmas are needed for some proofs in Chapter 4.

Theorem 2.1 (Chebyshev Alternation Theorem [30]). *Let D be a closed subset of $[a, b]$. Let $\{g_1, g_2, \cdots, g_n\}$ be a set of continuous functions on $[a, b]$ which satisfies the Haar's condition. A polynomial $p = \sum_i c_i g_i$ is the minimax approximate polynomial on D to any given continuous function f on D if and only if there are $n + 1$ elements $x_0 < \cdots < x_n$ in D such that $r(x_i) = -r(x_{i-1}) = \pm||r||_\infty$, $1 \leq i \leq n$ for the error function $r = f - p$.*

Remark 1. *Let D be $[-b, -a] \cup [a, b]$. Since $r(x_i) = \pm||r||_\infty$ for $0 \leq i \leq n$, $r(x)$ should have extreme points at x_i for $0 \leq i \leq n$. Thus, it holds that $p'(x_i) = 0$ and $x_i \in (-b, -a) \cup (a, b)$, or $x_i \in \{-b, -a, a, b\}$.*

Lemma 2.1 (Generalized de La Vallee Poussin Theorem [31]). *Let $\{g_1, g_2, \cdots, g_n\}$ be a set of continuous functions on $[a, b]$ that satisfies the Haar's condition. Let D be a closed subset of $[a, b]$ and let $f(x)$ be a continuous function on D. Let $x_i, 0 \leq i \leq n$ be $n+1$ consecutive points on D. Let $p(x)$ be a generalized polynomial such that $p - f$ has alternately positive and negative values at $x_i, 0 \leq i \leq n$. Let $p^*(x)$ be a minimax approximate polynomial on D for f and let $e(f)$ be the minimax approximaton error of $p^*(x)$. Then, it holds that*

$$e(f) \geq \min_i |p(x_i) - f(x_i)|.$$

Lemma 2.2 ([32]). *If $f(x)$ is an odd function, the minimax approximate polynomial of degree at most n to $f(x)$ is also odd function.*

2.7 Algorithms for Minimax Approximation

Remez algorithm [33] obtains the minimax approximate polynomials of a continuous function on one interval. It was proved that the Remez can always find the exact minimax approximate polynomials.

Recently, Lee et al. [31] proposed a modified Remez algorithm which finds the minimax approximate polynomial on multiple intervals and proved that the algorithm

can always find the minimax approximate polynomial for any piecewise continuous function. This modified Remez algorithm is used in this dissertation to find the minimax approximate polynomial for the sign function.

Let $\mu(x)$ be a function defined as

$$\mu(x) = \begin{cases} 1 & p(x) - f(x) \text{ is a local maximum value at } x \text{ on } D \\ -1 & p(x) - f(x) \text{ is a local minimum value at } x \text{ on } D \\ 0 & \text{otherwise.} \end{cases}$$

There are three criteria for choosing $n+1$ extreme points in Algorithm 5 as follows:

(i) Local extreme value condition; $\min_i \mu(y_i)(p(y_i) - f(y_i)) \geq E$.

(ii) Alternating condition; $\mu(y_i) \cdot \mu(y_{i+1}) = -1$ for $i = 1, \cdots, n$.

(iii) Maximum absolute sum condition; $\sum_{i=1}^{n+1} |p(y_i) - f(y_i)|$ is maximum for all candidate set of extreme points satisfying the local extreme value condition and the alternating condition.

The modified Remez algorithm operates with n basis functions $\{g_1, g_2, \cdots, g_n\}$. Suppose that the minimax approximate polynomial $p(x)$ is represented with the basis functions as $p(x) = \sum_{i=1}^{n} c_i g_i(x)$. The modified Remez algorithm finds the coefficients c_i's of $p(x)$. The simplest basis functions are a power basis, $\{1, x, x^2, \cdots, x^{n-1}\}$. However, when approximating the sign function using this basis, the magnitudes of the coefficients c_i's are unstable such as too small values or too large values, which makes a lot of numerical errors. Therefore, the Chebyshev polynomials are usually used as the basis functions. The Chebyshev polynomials T_i's on $[-1, 1]$ are defined by the following recursion;

$$T_0(t) = 1$$

$$T_1(t) = t$$

$$T_i(t) = 2tT_{i-1}(t) - T_{i-2}(t) \text{ for } i \geq 2.$$

If the sign function is approximated on a domain $[-b, b]$ for some $b > 1$, then $\tilde{T_i}(t) = T_i(t/b)$ should be used instead of T_i for all i.

Algorithm 4: Remez algorithm [31]

Input: Polynomial basis $\{g_1, \cdots, g_n\}$, a domain $[a, b]$, an approximation parameter δ, and a continuous function f on $[a, b]$

Output: The minimax approximate polynomial p for f

1 Choose $x_1, \cdots, x_{n+1} \in [a, b]$, where $x_1 < \cdots < x_{n+1}$;

2 Find the polynomial $p(x)$ in terms of $\{g_1, \cdots, g_n\}$ such that $p(x_i) - f(x_i) = (-1)^i E, 1 \leq i \leq n+1$ for some E;

3 Divide the domain $[a, b]$ into $n+1$ sections $[z_{i-1}, z_i], i = 1, \cdots, n+1$. $z_1, \cdots, z_n$ are zeros of $p(x) - f(x)$, where $x_i < z_i < x_{i+1}$, and $z_0 = a, z_{n+1} = b$;

4 Find the maximum or minimum point for each section when $p(x_i) - f(x_i)$ has positive or negative value, respectively. These points $y_1, \cdots, y_{n+1}$ are called extreme points;

5 $\epsilon_{max} \leftarrow \max_{1 \leq i \leq n+1} |p(y_i) - f(y_i)|$;

6 $\epsilon_{min} \leftarrow \min_{1 \leq i \leq n+1} |p(y_i) - f(y_i)|$;

7 **if** $(\epsilon_{max} - \epsilon_{min})/\epsilon_{min} < \delta$ **then**

8 Return $p(x)$;

9 **else**

10 Replace x_i's with y_i's. Go to line 2;

11 **end**

Algorithm 5: Modified Remez algorithm [31]

Input: A polynomial basis $\{g_1, \cdots, g_n\}$, an approximation parameter δ, an input domain $D = \bigcup_{i=1}^{l}[a_i, b_i] \subset \mathbb{R}$, and a continuous function f on D

Output: The minimax approximate polynomial p for f

1 Choose $x_1, \cdots, x_{n+1} \in D$, where $x_1 < \cdots < x_{n+1}$;

2 Find the polynomial $p(x)$ in terms of $\{g_1, \cdots, g_n\}$ such that $p(x_i) - f(x_i) = (-1)^i E, 1 \leq i \leq n+1$ for some E;

3 Collect all the extreme and boundary points such that $\mu(x)(p(x) - f(x)) \geq |E|$ and put them in a set B;

4 Find $n+1$ extreme points $y_1 < y_2 < \cdots < y_{n+1}$ in B which satisfy alternating condition and maximum absolute sum condition;

5 $\epsilon_{max} \leftarrow \max_{1 \leq i \leq n+1} |p(y_i) - f(y_i)|$;

6 $\epsilon_{min} \leftarrow \min_{1 \leq i \leq n+1} |p(y_i) - f(y_i)|$;

7 **if** $(\epsilon_{max} - \epsilon_{min})/\epsilon_{min} < \delta$ **then**

8 Return $p(x)$;

9 **else**

10 Replace x_i's with y_i's. Go to line 2;

11 **end**

Chapter 3

Improvement of FrodoKEM Using Gray and BCH Codes

Existing public key cryptosystems such as RSA and elliptic curve cryptography can be broken by future quantum computers because of the rapid development of quantum computers. Therefore, it is very important to develop secure PQC algorithms resistant to quantum computing. Currently, NIST is in the process of proposing, evaluating, and standardizing PQC algorithms. In the NIST PQC Round 1, the 26 algorithms have been selected for NIST PQC Round 2 and are under evaluation. Notably, 12 of the 26 algorithms are lattice-based ones. Thus, lattice-based cryptography is clearly the most promising field for PQC [4, 5].

LWE is a problem presented by Regev [6] in 2005 and is reduced to worst-case problems on lattices. Ring-LWE (RLWE) is a problem presented in [34], where it is reduced to worst-case problems on ideal lattices. RLWE significantly reduces the key size of cryptosystems based on LWE. Many lattice-based PKE and KEM schemes submitted to NIST are based on the hardness of LWE and RLWE.

Among the lattice-based algorithms selected for NIST PQC Round 2, many proposed algorithms use ECCs to improve their performances. NewHope [2] uses a simple error correction technique ATE. Round5 [8], which is a combined algorithm of Hila5

[9] and Round2 [10], uses an ECC called XE5 which is resistant to side-channel attacks. LAC [11, 12] uses BCH codes of large code lengths. Three Bears [13] uses BCH codes such that constant-time implementation is possible, but its performance is relatively worse compared to those of the BCH codes used in LAC. A lattice-based KEM scheme called KCL [14], which was submitted to NIST PQC Round 1 but was not selected for NIST PQC Round 2, uses a single-error correcting code, lattice code in $\tilde{D}_4$ [15], or lattice code in E_8 [14]. However, all of these lattice-based algorithms using ECCs are ring-based schemes, and there is no case of using ECCs for non-ring ones such as FrodoKEM [1, 35, 36]. Here, I want to emphasize that my research is the first-ever one to apply and analyze both ECCs and Gray coding to a non-ring lattice-based KEM, FrodoKEM. In addition, the application of ECCs and Gray coding to non-ring schemes such as FrodoKEM is not straightforward, and hence, various new ideas have been applied to the proposed results as explained here.

FrodoKEM is one of the representative PQC schemes selected for the NIST PQC Round 2. Therefore, improving the performance of FrodoKEM is considerably important, and the proposed schemes in this dissertation are possibly applied to other non-ring schemes. In this dissertation, I aim to improve the performance of FrodoKEM as follows:

(i) There is a risk that PQC cryptosystems will be broken because of the increasing computing power in the era of quantum computers. Therefore, I am motivated to work on how to use ECCs to improve the security level of FrodoKEM so that FrodoKEM can resist enhanced computing power in the coming future.

(ii) In the IoT era, it is very important for cryptosystems to be able to send multiple keys simultaneously or to reduce the bandwidth. Therefore, I work on how to use ECCs to increase the message size so that multiple keys can be simultaneously sent and to reduce the bandwidth.

In this chapter, how to apply ECCs to FrodoKEM [1] is studied to improve its secu-

rity level and/or lower its bandwidth. I propose a method to apply ECCs to FrodoKEM by encoding the bits converted from the encrypted symbols. In addition, the DFR is reduced using Gray codes as bit-to-symbol mapping. The proposed method has the advantage of improving performances without modifying the existing framework of FrodoKEM. Note that the combination of ECCs and Gray coding is widely used in the field of wireless communication systems to lower the bit error probability in higher-order modulation such as PAM and QAM [37]. However, although such a combination has been widely used, I, for the first time, apply it to a lattice-based PKE or KEM scheme. The symbols of $\mathbb{Z}_q$ considered in this chapter are similar to those of PAM, but modulo q operations should be performed after adding errors. It is called modulo q PAM. In addition, the environment of wireless communication systems is quite different from that of lattice-based PKE / KEM schemes as follows. In the case of PAM for wireless communications, when an error is added to a symbol with the largest magnitude, it is saturated rather than changed to another symbol. However, for the lattice-based PKE / KEM schemes, all the symbols of $\mathbb{Z}_q$ are computed by modulo q operations. For example, if an error 1 is added to the largest value $q-1$, it becomes the smallest value 0.

The limited-magnitude error control codes [38] often adopt Gray coding. However, while the errors in the channel model for those codes are asymmetric and limited in their magnitude, the errors in the channel model for FrodoKEM are symmetric and not limited in their magnitude under mod q arithmetic.

In [16], the performance of NewHope was improved by using ECCs. The ATE technique used in NewHope is replaced by BCH codes or concatenated coding schemes of low-density parity check (LDPC) and BCH codes to improve the security level. However, the application of ECCs to FrodoKEM is quite different from that in [16] for the following reasons:

(i) The original FrodoKEM does not use ECCs, and thus, ECCs should be carefully applied to FrodoKEM. To properly apply ECCs to FrodoKEM, it is needed to

change some parameters of FrodoKEM, and thus, there is no guarantee that the performance will be improved as much as expected even if ECCs are used. Its performance can be improved because I carefully select ECCs, change the parameter values, and use Gray coding with modulo q PAM.

(ii) Unlike the schemes based on ring-LWE such as NewHope, ECCs cannot be easily applied to non-ring schemes such as FrodoKEM because the number of symbols to which message bits is mapped to be very small. For example, NewHope and FrodoKEM have 1024 symbols and 64 symbols, respectively. Thus, it is not easy to design effective ECCs for the small number of message symbols in FrodoKEM. Furthermore, because multiple bits are mapped to one symbol in FrodoKEM, an error of one symbol can result in more than one error bit. It is complicated to calculate the DFR by considering more than one error bit. However, it is shown that the probability that errors to the non-adjacent symbols occur is relatively negligible and thus Gray coding is essential. Thus, by applying the Gray code in the ECCs, one symbol error can be regarded as one error bit, which makes it possible to calculate the DFR.

(iii) Assume that normal bit-to-symbol mapping is a mapping in which both bit string and symbol size are in an increasing order. The mappings in (2.1), (2.2), and (3.5) in this dissertation are normal bit-to-symbol mappings. In contrast, the Gray mappings in (3.1) and (3.2) are not normal bit-to-symbol mappings. The bit strings are not in increasing order in these Gray mappings. When using normal bit-to-symbol mappings without Gray coding, the DFR becomes very high because more error bits occur with higher probability, and thus, the performance of FrodoKEM is not improved, as given in the Tables 3.2, 3.4, and 3.5.

I propose and analyze combined schemes of ECCs and Gray coding in a non-ring scheme, FrodoKEM for the first time; thus, the performance of FrodoKEM is substantially improved. My contribution can be summarized as follows:

(i) The security level of FrodoKEM is improved by increasing the standard deviation σ of error distribution. Because the DFR increases as σ increases, the DFR requirement is satisfied by properly using Gray and BCH codes with modulo q PAM.

(ii) The number of message bits of Frodo-640 is increased from 128 bits to 192 bits while keeping the required security level. If the number of message bits increases, σ should be reduced to meet the DFR requirement, which leads to degradation of the security level. Such security level degradation can be avoided by properly using Gray and BCH codes with modulo q PAM.

(iii) The bandwidth of FrodoKEM is reduced by using a smaller q. Because the DFR increases as q decreases, the DFR requirement is satisfied by properly using Gray and BCH codes with modulo q PAM.

3.1 Modification of FrodoKEM with Gray and Error-Correcting Codes

3.1.1 Viewing FrodoPKE as a Digital Communication System

To apply ECCs to FrodoPKE and analyze them, it is convenient to understand the FrodoPKE in terms of digital communication systems, where messages are transmitted to the receiver via an encoder, modulator, (noisy) channel, demodulator, and decoder. Figure 3.1 shows the description of Frodo-640 as a digital communication system.

In this model, the sender is Alice and the receiver is Bob. The shared key μ that Alice wants to send corresponds to the message bits. The mapping of binary bits to symbols in $\mathbb{Z}_q$ in FrodoPKE corresponds to modulation. $Frodo.Encode$ function uses the term 'encode', but in fact it corresponds to a modulator in the digital communication. In this chapter, $Frodo.Encode$ in FrodoPKE is referred to as a modulator.

In FrodoPKE as in Algorithms 1, 2, and 3, Alice computes $Frodo.Encode(\mu)$ to

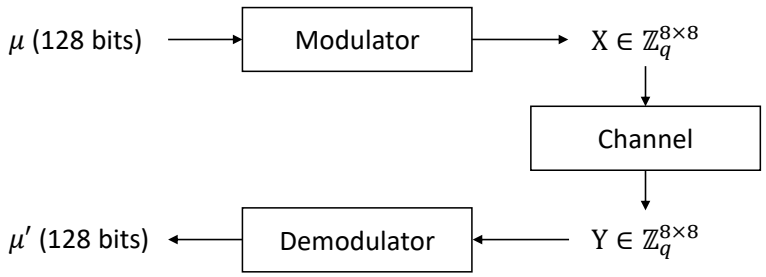

Figure 3.1: Description of Frodo-640 as a digital communication system.

generate two ciphertexts $\mathbf{C_1}$ and $\mathbf{C_2}$ and sends them. Bob computes $\mathbf{M} = \mathbf{C_2} - \mathbf{C_1}\mathbf{S}$ with the received $\mathbf{C_1}$, $\mathbf{C_2}$ and the secret key $\mathbf{S}$. As a result, $Frodo.Encode(\mu)$ is added with noise $\mathbf{E'''}$. This procedure can be seen as $Frodo.Encode(\mu)$ passing through a noisy channel in the digital communication. Here, the noise element of $\mathbf{E'''}$ follows χ' described in Section 2.3.2. These noise elements are not i.i.d. However, because exact analysis is difficult, it is assumed that they are i.i.d.

The $Frodo.Decode$ function works as follows. The $Frodo.Decode$ function corresponds to a demodulator in the digital communication. $Frodo.Decode(\mathbf{M})$ rounds each symbol in $\mathbb{Z}_q$ of the received matrix with errors to the nearest multiple of $q/4$ or $q/8$ for Frodo-640 and Frodo-976, respectively. Then, the inverse of the mapping in (2.1) or (2.2) is applied to obtain the estimated bit string μ'.

If ECCs are used in FrodoPKE, encoding is added before modulation, and decoding is added after demodulation. Figure 3.2 shows the application of ECC to Frodo-640 as a digital communication system.

3.1.2 Error-Correcting Codes for FrodoPKE

To meet security categories 1 and 3 in NIST PQC Standardization, the obtained DFR should be less than 2^{-128} and 2^{-192}, respectively. Among various ECCs, algebraic codes are used, especially BCH codes, rather than modern codes such as LDPC codes for the following reasons:

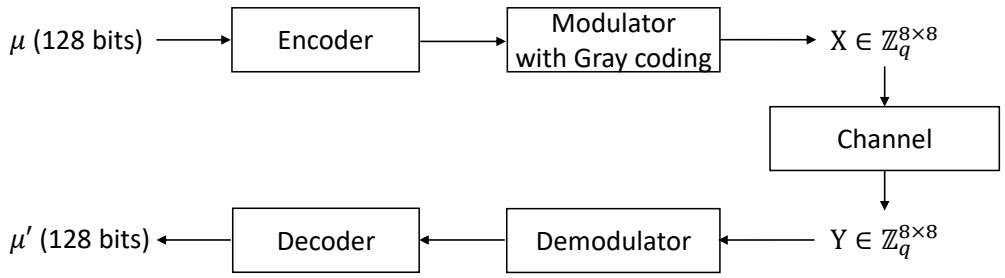

Figure 3.2: Description of Frodo-640 with ECC as a digital communication system.

(i) LDPC codes have a serious error floor problem to be used for FrodoKEM. Because the error floor is reached quickly, the performance of LDPC codes is much worse than that of the algebraic codes for the region of DFR lower than 2^{-128}.

(ii) For LDPC codes, it is difficult to algebraically calculate the DFR, and thus, the DFR should be estimated through numerical analysis. However, in FrodoKEM, the DFR should be less than 2^{-128} or 2^{-192}, and numerical analysis in this error range is impossible. They [16] could not calculate the DFR for LDPC code for this low DFR.

(iii) For the concatenated coding schemes of LDPC and algebraic codes as in [16], it is needed to know the statistical characteristics of errors remaining after LDPC decoding to algebraically estimate the DFR. Although it was not clearly stated in [16], they seem to assume that the errors remaining after LDPC decoding are statistically independent and uniformly distributed. However, the LDPC decoding errors tend to be bursty, and the analysis of characteristics of LDPC decoding errors is known to be a hard problem in the field of coding theory. In addition, the block error rate around 2^{-128} is the range where numerical analysis is impossible.

Thus, in this dissertation, BCH codes are used because they provide various parameter values. Specifically, binary BCH codes with parameters (192, 128, 8), (256,

128, 18), (256, 192, 8), and (256, 192, 8) are used. These are modified BCH codes obtained by shortening [39] or extending [40] the original BCH codes. For a systematic BCH code, the error correcting capability l_t is not reduced from shortening. With extension, l_t is maintained or increased. Using these properties, these BCH code parameters are found to be suitable for FrodoKEM. Note that shortening and extending do not significantly affect the encoding and decoding algorithms and complexity.

3.1.3 Gray Coding

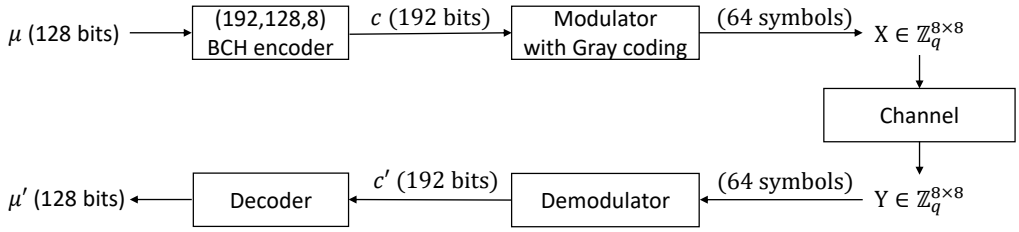

Figure 3.3: Frodo-640 with (192, 128, 8) BCH and Gray codes.

It is well known that Gray coding should be used to map binary data to symbols from large alphabet for better bit error correction performance in digital communication. For example, consider the case of applying the (192, 128, 8) BCH code to Frodo-640 as in Figure 3.3. Encoding the 128-bit message μ results in a 192-bit codeword c. In the modulation with $B = 3$, each of the three bits in the codeword c is mapped to a symbol in $\mathbb{Z}_q$ according to the following Gray coding, which is different from the mapping in (2.2) as:

- $B = 3$ with Gray coding;

$$\begin{aligned} &000 \to 0, \quad 001 \to \frac{q}{8}, \quad 011 \to \frac{2q}{8}, \quad 010 \to \frac{3q}{8}, \\ &110 \to \frac{4q}{8}, \quad 111 \to \frac{5q}{8}, \quad 101 \to \frac{6q}{8}, \quad 100 \to \frac{7q}{8}. \end{aligned} \tag{3.1}$$

Gray coding in (3.1) is depicted in Figure 3.4, where dotted lines denote decision

boundaries for demodulation. Figure 3.4 shows that the bit difference between adjacent symbols is always one bit in Gray coding. The reason for using Gray coding is to minimize the number of bit errors and increase the error correction probability of ECCs.

If ECCs are not used, one symbol error immediately causes a decryption failure. Thus, it is meaningless to minimize the number of bit errors by using Gray coding for one symbol error. However, when ECCs are used, the total bit errors whose number does not exceed error correction capacity are correctable. Thus, it is important to reduce the number of bit errors by using Gray coding.

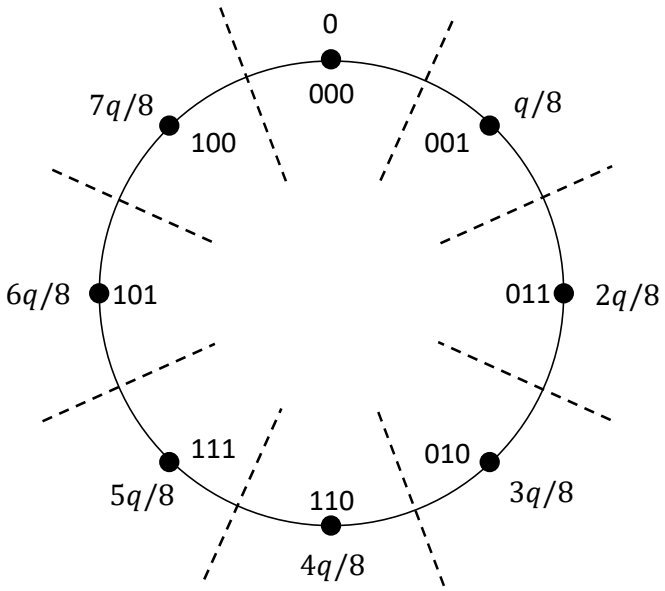

Figure 3.4: Gray coding for $B = 3$.

The 192-bit codeword is mapped to 64 symbols in $\mathbb{Z}_q$, and errors are added to these symbols in $\mathbb{Z}_q$ while passing through the channel. Demodulation rounds each symbol in $\mathbb{Z}_q$ to the nearest multiples of $q/8$ and then applies the inverse of the mapping in (3.1). Then, c' is obtained, which is the codeword c added with errors. Then, BCH decoding is performed to obtain c to estimate μ.

Gray coding for $B = 4$ is performed as:

- $B = 4$ with Gray coding;

$$\begin{aligned}
&0000 \to 0, \quad 0001 \to \frac{q}{16}, \quad 0011 \to \frac{2q}{16}, \quad 0010 \to \frac{3q}{16}, \\
&0110 \to \frac{4q}{16}, \quad 0111 \to \frac{5q}{16}, \quad 0101 \to \frac{6q}{16}, \quad 0100 \to \frac{7q}{16}, \\
&1100 \to \frac{8q}{16}, \quad 1101 \to \frac{9q}{16}, \quad 1111 \to \frac{10q}{16}, \quad 1110 \to \frac{11q}{16}, \\
&1010 \to \frac{12q}{16}, \quad 1011 \to \frac{13q}{16}, \quad 1001 \to \frac{14q}{16}, \quad 1000 \to \frac{15q}{16}.
\end{aligned} \tag{3.2}$$

Figure 3.5 describe applying the (256, 128, 18) BCH code to Frodo-640.

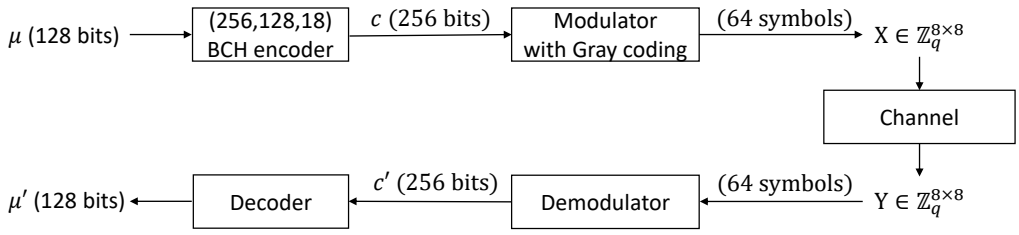

Figure 3.5: Frodo-640 with (256, 128, 18) BCH and Gray codes.

3.1.4 IND-CCA Security of Modified FrodoKEM

ECCs are applied to FrodoPKE as shown in Figure 3.2. BCH encoder $BCH.Encode$ is added before $Frodo.Encode$ which corresponds to modulator and BCH decoder $BCH.Decode$ is added after $Frodo.Decode$ which corresponds to demodulator. Specific BCH parameters are shown in Section 3.1.2. In addition, Gray coding in (3.1) or (3.2) is used instead of the existing $Frodo.Encode$ function, and $Frodo.Decode$ is replaced with another function as shown in Section 3.1.3. Let us call the modified $Frodo.Encode$ as $Frodo.Encode'$ and the modified FrodoPKE as FrodoPKE′.

In this dissertation, I describe FrodoPKE′ and analyze the performance of the proposed modified FrodoKEM scheme through the analysis of FrodoPKE′ without

the description of the proposed KEM scheme for simplicity. The proposed modified FrodoKEM is derived from FrodoPKE′ using QFO transformation, similar to that in the previous study [1]. The description of the proposed KEM scheme can be omitted for the following two reasons. First, the QFO transformation method described in [1] can be used almost identically to construct the proposed modified KEM. In addition, it is possible to analyze the performance of the proposed modified FrodoKEM according to the change in parameters only by describing the FrodoPKE′.

Consider the IND-CCA security of the proposed modified KEM scheme. The IND-CCA security proof of FrodoKEM shown in [1] is summarized as follows. They proved that FrodoPKE achieves IND-CPA security and then proved that FrodoKEM modified using QFO transformation achieves IND-CCA security. Similarly, if FrodoPKE′ is proved to achieve IND-CPA security, then the proposed modified FrodoKEM can also achieve IND-CCA security. The proposition that FrodoPKE′ has IND-CPA security can be proved by the fact that FrodoPKE achieves IND-CPA security as follows.

(i) Suppose that there is an algorithm $\mathcal{A}$ that attacks FrodoPKE′. Let us design an algorithm $\mathcal{A}'$ that uses $\mathcal{A}$ to attack FrodoPKE.

(ii) $\mathcal{A}$ claims to be able to distinguish between ciphertexts of μ_0 and μ_1.

(iii) $\mathcal{A}'$ puts the following values into FrodoPKE as inputs; $Frodo.Encode^{-1}[Frodo.Encode'(BCH.Encode(\mu_i))]$ for $i = 0$ and 1. $\mathcal{A}'$ receives the ciphertexts ct_0 and ct_1. Then, ct_0 and ct_1 correspond to ciphertexts of μ_0 and μ_1 in FrodoPKE′.

(iv) $\mathcal{A}'$ passes the received ct_0 and ct_1 to $\mathcal{A}$ to distinguish.

(v) $\mathcal{A}$ distinguishes ct_0 and ct_1, and computes $b \in \{0, 1\}$.

(vi) $\mathcal{A}'$ distinguishes the ciphertexts of the original messages μ_0 and μ_1 by outputting b calculated by $\mathcal{A}$.

Since the proposed scheme simply adds BCH encoding and BCH decoding algorithms to FrodoKEM, there is no danger of being particularly vulnerable to primal and dual attacks, which are simply LWE attacks.

3.1.5 Evaluation of DFR

In FrodoKEM, 64 symbols in $\mathbb{Z}_q$ are transmitted. Given the parameters n, q, σ, B, and the maximum number of correctable bit errors t using an ECC (t = 0 if ECC is not used), the DFR is computed through the following procedures:

(i) Find an optimal discrete noise distribution χ that approximates the rounded continuous Gaussian for a given σ. Find the optimal distribution considering attack time for the LWE problem.

(ii) Obtain the product distribution ψ of two optimal distributions χ. Convolve ψ $2n$ times and then convolve the resulting distribution with χ. Then, a distribution χ' is obtained.

(iii) Compute the symbol error rate (SER) p with χ' as:

$$\mathrm{SER} = \sum_{x \in [q/2^{B+1}, q-q/2^{B+1})} \chi'(x). \tag{3.3}$$

(iv) Compute the DFR of FrodoKEM using an ECC as follows. When an ECC with the error correction capability t is used, Gray coding should be used. The 64 B-bit messages are encoded into 64 symbols in $\mathbb{Z}_q$, with the probability p that an error occurs for each symbol. When Gray coding is used, most symbol errors only generate one-bit errors. For example, Figure 3.4 shows that only a one-bit error occurs when the decision boundary is crossed once. For various cases that will be described in Section 3.2, Table 3.1 shows the comparison between the probability of crossing the decision boundary once and that of crossing the decision boundary twice. The probability of crossing the decision boundary twice is relatively negligible, as shown in Table 3.1, and in Case 5), it is too small; thus,

the value cannot be obtained. The DFR is the probability that more than t symbol errors out of 64 symbols occur, which can be obtained using the following equation

$$\mathrm{DFR} = \sum_{i=t+1}^{64} \binom{64}{i} p^i (1-p)^{64-i}. \tag{3.4}$$

Table 3.1: Comparison between the probability of crossing the decision boundary once and that of crossing the decision boundary twice.

	σ	B	probability of crossing decision boundary once	probability of crossing decision boundary twice
Case 1)	3.38	3	$2^{-20.42}$	$2^{-72.73}$
Case 2)	2.87	4	$2^{-10.72}$	$2^{-36.73}$
Case 3)	2.84	4	$2^{-26.06}$	$2^{-94.52}$
Case 4)	1.93	3	$2^{-155.62}$	$2^{-528.52}$
Case 5)	1.78	3	$2^{-203.53}$	.
Case 6)	2.38	4	$2^{-20.45}$	$2^{-72.87}$
Case 7)	2.22	4	$2^{-26.15}$	$2^{-93.92}$
Case 8)	2.3	2	$2^{-82.42}$	$2^{-291.49}$
Case 9)	2.3	3	$2^{-23.04}$	$2^{-82.41}$
Case 10)	2.3	3	$2^{-7.00}$	$2^{-23.04}$

When using an ECC, if Gray coding is not used, the mapping in (2.2) for $B = 3$ and the following mapping for $B = 4$ can be used:

- $B = 4$ without Gray coding;

$$
\begin{aligned}
&0000 \to 0, \quad 0001 \to \frac{q}{16}, \quad 0010 \to \frac{2q}{16}, \quad 0011 \to \frac{3q}{16}, \\
&0100 \to \frac{4q}{16}, \quad 0101 \to \frac{5q}{16}, \quad 0110 \to \frac{6q}{16}, \quad 0111 \to \frac{7q}{16}, \\
&1000 \to \frac{8q}{16}, \quad 1001 \to \frac{9q}{16}, \quad 1010 \to \frac{10q}{16}, \quad 1011 \to \frac{11q}{16}, \\
&1100 \to \frac{12q}{16}, \quad 1101 \to \frac{13q}{16}, \quad 1110 \to \frac{14q}{16}, \quad 1111 \to \frac{15q}{16}.
\end{aligned} \tag{3.5}
$$

Mappings in (2.1) and (2.2) are given in FrodoKEM. However, there is no bit-to-symbol mapping for $B = 4$ in FrodoKEM; thus, the bit-to-symbol mapping in (3.5) is given for $B = 4$ in the same way as the mappings in (2.1) and (2.2).

In this dissertation, the DFR of FrodoKEM using an ECC with and without Gray coding is computed and compared. The following procedure depicts how to compute the DFR of FrodoKEM using an ECC without Gray coding. Let p_1, p_2, p_3, and p_4 be the probability values that the number of bit errors in one symbol in $\mathbb{Z}_q$ after demodulation is 1, 2, 3, and 4, respectively. Let n_i, $i \in 1, 2, 3, 4$, be the number of symbols in $\mathbb{Z}_q$ such that the number of bit errors in a symbol is i. Suppose that the codeword bits are uniform at random. Then, p_1, p_2, and p_3 are approximately $p/2$, $p/4$, and $p/4$ when $B = 3$, respectively. In addition, p_1, p_2, p_3, and p_4 are approximately $p/2$, $p/4$, $p/8$, and $p/8$ when $B = 4$, respectively. Then, the DFRs of FrodoKEM using ECCs without Gray coding are obtained for $B = 3$ and $B = 4$, respectively, as follows;

$$
\sum_{n_1+2n_2+3n_3>t} \binom{64}{n_1, n_2, n_3, 64-n_1-n_2-n_3} {p_1}^{n_1} {p_2}^{n_2} {p_3}^{n_3} (1-p)^{64-n_1-n_2-n_3} \tag{3.6}
$$

$$\sum_{n_1+2n_2+3n_3+4n_4>t} \binom{64}{n_1, n_2, n_3, n_4, 64-n_1-n_2-n_3-n_4} {p_1}^{n_1}{p_2}^{n_2}{p_3}^{n_3}{p_4}^{n_4}(1-p)^{64-n_1-n_2-n_3-n_4}. \tag{3.7}$$

3.1.6 Error Dependency

Because it is very difficult to derive the DFR considering error dependency, the DFR is usually computed assuming that errors are statistically independent. Recent studies have reported that the DFR was underestimated when using ECC [41] because error dependency is not considered, and they proposed a DFR calculation method for LAC considering error dependency. However, their DFR calculation method cannot be applied to other schemes including FrodoKEM. The DFR is computed based on the assumption that the error coefficients of $\mathbf{S}'\mathbf{E} + \mathbf{E}'' - \mathbf{E}'\mathbf{S}$ in Algorithm 3 are independent. Considering the DFR deviation due to the independence assumption, I try to set a margin that is enough in the DFRs of the proposed improvements. The exact analysis of DFR considering error dependency will be studied in future work.

3.2 Performance Improvement of FrodoKEM Using Gray and BCH Codes

As ECCs, BCH codes are used to improve the security levels, increase the message size, and reduce the bandwidth of Frodo-640 and Frodo-976.

3.2.1 Improving the Security Level of FrodoKEM

Because B should be increased to use ECCs, the DFR also increases. However, using BCH codes, the DFR can be effectively lowered, and thus the security level is improved by maximizing σ while satisfying DFR $< 2^{-148.8}$ for Frodo-640 and DFR

Table 3.2: Cases for improving the security level of FrodoKEM scheme.

	message size [bits]	σ	security level [bits]	SER with Gray coding	DFR with Gray coding	SER without Gray coding	DFR without Gray coding
Frodo-640	128	2.75	149.30	$2^{-154.82}$	$2^{-148.82}$	.	.
Frodo-976	192	2.3	215.66	$2^{-205.56}$	$2^{-199.56}$	.	.
Case 1)	128	3.38	156.98	$2^{-20.42}$	$2^{-149.05}$	$2^{-20.42}$	$2^{-51.90}$
Case 2)	128	2.87	152.25	$2^{-10.72}$	$2^{-150.71}$	$2^{-10.72}$	$2^{-50.77}$
Case 3)	192	2.84	225.97	$2^{-26.06}$	$2^{-199.88}$	$2^{-26.06}$	$2^{-68.52}$

$< 2^{-199.6}$ for Frodo-976. The parameters $n, q, \bar{n}$, and $\bar{m}$ are maintained, but only σ and B are adjusted. Then, the security level can be improved while maintaining the bandwidth and satisfying the DFR requirement. Table 3.2 summarizes the performances of various cases with and without Gray coding to improve the security level as explained below.

Case 1) Frodo-640 with (192, 128, 8) BCH code:

Encode the 128-bit message with (192, 128, 8) BCH code to obtain a 192-bit codeword c. Modulates c using Gray coding in (3.1), and then, c passes through the channel. From Table 3.2, B, σ, and the security level are changed as;

- B; $2 \rightarrow 3$
- σ; $2.75 \rightarrow 3.38$
- security level; 149.30 $\rightarrow$ 156.98
- SER = $2^{-20.42}$
- DFR = error probability after BCH decoding = probability of more than 8 errors = $2^{-149.05}$.

If the mapping in (2.2) instead of Gray coding in (3.1) is used, SER and DFR are calculated as

- SER = $2^{-20.42}$
- DFR = $2^{-51.90}$.

From these results, it can be seen that the security level is increased by applying the BCH code to Frodo-640. In addition, Gray coding is essential because when Gray coding is not used, the DFR is much higher than the DFR when Gray coding is used.

Note that the security level can also be improved by simply increasing n without using the BCH code. However, the bandwidth also increases as n increases. Table 3.3 compares Frodo-640 with (192, 128, 8) BCH code and Frodo-640 with increased n while all parameters other than B and n (dimension of matrices) are unchanged. In Frodo-640 with increased n, the security level is improved, but the bandwidth also increases. Case 1) improves the security level while maintaining the bandwidth.

Table 3.3: Comparison of Frodo-640 with (192, 128, 8) BCH code and Frodo-640 with increased n.

	security level [bits]	DFR	public key size [bytes]	ciphertext size [bytes]
Frodo-640 with (192, 128, 8) BCH code	149 $\rightarrow$ 157	$2^{-149.05}$	9616	9736
Frodo-640 $n : 640 \rightarrow 700$	149 $\rightarrow$ 166	$2^{-137.31}$	10516	10636

Case 2) Frodo-640 with (256, 128, 18) BCH code:

Encode the 128-bit message with the (256, 128, 18) BCH code to obtain a 256-bit codeword c. Modulate c using Gray coding in (3.2), and then, c passes through the

channel. From Table 3.2, B, σ, and the security level are changed as;

- B; $2 \rightarrow 4$
- σ; $2.75 \rightarrow 2.87$
- security level; 149.30 $\rightarrow$ 152.25
- SER = $2^{-10.72}$
- DFR = error probability after BCH decoding =
  probability of more than 18 errors = $2^{-150.71}$.

If the mapping in (3.5) instead of Gray coding in (3.2) is used, SER and DFR can be calculated as

- SER = $2^{-10.72}$
- DFR = $2^{-50.77}$.

From these results, it can be seen that the security level is increased by applying the BCH code to Frodo-640. However, the security level of Case 2) is less than the security level of Case 1). In addition, Gray coding is essential similar to Case 1).

Case 3) Frodo-976 with (256, 192, 8) BCH code:

Encode the 192-bit message with the (256, 192, 8) BCH code to obtain a 256-bit codeword c. Modulate c using Gray coding in (3.2), and then, c passes through the channel. From Table 3.2, B, σ, and the security level are changed as;

- B; $3 \rightarrow 4$
- σ; $2.3 \rightarrow 2.84$
- security level; 215.66$\rightarrow$225.97
- SER = $2^{-26.06}$
- DFR = error probability after BCH decoding =
  probability of more than 8 errors = $2^{-199.88}$.

If the mapping in (3.5) instead of Gray coding in (3.2) is used, SER and DFR are calculated as

- SER = $2^{-26.06}$
- DFR = $2^{-68.52}$.

From these results, it can be seen that the security level is increased by applying the BCH code to Frodo-976. In addition, Gray coding is essentially similar to the previous cases.

3.2.2 Increasing the Message Size of Frodo-640

Table 3.4: Cases for increasing the message size of the FrodoKEM scheme.

	message size [bits]	σ	security level [bits]	SER with Gray coding	DFR with Gray coding	SER without Gray coding	DFR without Gray coding
Frodo-640	128	2.75	149.30	$2^{-154.82}$	$2^{-148.82}$	·	·
Frodo-976	192	2.3	215.66	$2^{-205.56}$	$2^{-199.56}$	·	·
Case 4)	192	1.93	137.18	$2^{-155.62}$	$2^{-149.62}$	·	·
Case 5)	192	1.78	134.52	$2^{-207.62}$	$2^{-201.62}$	·	·
Case 6)	192	2.38	144.27	$2^{-20.45}$	$2^{-149.35}$	$2^{-20.45}$	$2^{-51.68}$
Case 7)	192	2.22	141.91	$2^{-26.15}$	$2^{-200.70}$	$2^{-26.15}$	$2^{-68.79}$

In this section, the 192-bit message for Frodo-640 instead of the 128-bit message is used because increasing the message bits from 128 bits to 192 bits has several advantages. The 128-bit key for symmetric key encryption and 64-bit key for authentication can be sent at the same time via the 192-bit key. In addition, if the 80-bit lightweight cryptographic keys for IoT systems are used, two 80-bit keys can be sent at once.

In the following Cases 5) and 6), 256-bit codewords are mapped to the transmitted

matrix in $\mathbb{Z}_q^{8\times 8}$, and thus, B is increased from 2 to 4. Because of the increase in B, the DFR also increases, and σ should be decreased to satisfy the DFR requirement. Then, the security level significantly decreases. However, it is possible to prevent degradation of the security level by using BCH codes. Case 4) uses the 192-bit message without using the BCH code. However, the security level significantly decreases. Cases 5) and 6) using BCH codes can satisfy DFR $< 2^{-148.82}$ and DFR $< 2^{-199.56}$, respectively. Table 3.4 summarizes the performances of various cases with and without Gray coding to increase the message as explained below.

Case 4) Frodo-640, Message; 128 $\rightarrow$192 bits, DFR $< 2^{-148.82}$:

The 192-bit message μ is modulated using the mapping in (2.2), and then, it passes through the channel. Then, from Table 3.4, B, σ, and the security level are changed as;

- B; $2 \rightarrow 3$
- σ; $2.3 \rightarrow 1.93$
- security level; 149.30 $\rightarrow$ 137.18
- SER = $2^{-155.62}$
- DFR = $2^{-149.62}$.

It can seen that increasing the message size decreases the security level significantly.

Case 5) Frodo-640, Message; 128 $\rightarrow$192 bits, DFR $< 2^{-199.56}$:

The 192-bit message μ is modulated using the mapping in (2.2), and then, it passes through the channel. Then, from Table 3.4, B, σ, and the security level are changed as;

- B; $2 \rightarrow 3$
- σ; $2.3 \rightarrow 1.78$

- security level; 149.30 $\rightarrow$ 134.52
- SER = $2^{-207.62}$
- DFR = $2^{-201.62}$.

It can be seen that increasing the message size decreases the security level significantly.

Case 6) Frodo-640 with (256, 192, 8) BCH code, Message; 128 $\rightarrow$192 bits, DFR $< 2^{-148.82}$:

Encode the 192-bit message with the (256, 192, 8) BCH code to obtain the 256-bit codeword c. c is modulated using Gray coding in (3.2), and it passes through the channel. From Table 3.4, B, σ, and the security level are changed as;

- B; $2 \rightarrow 4$
- σ; $2.75 \rightarrow 2.38$
- security level; 149.30 $\rightarrow$ 144.27
- SER = $2^{-20.45}$
- DFR = error probability after BCH decoding =
  probability of more than 8 errors = $2^{-149.35}$.

If the mapping in (3.5) instead of Gray coding in (3.2) is used, SER and DFR are calculated as

- SER = $2^{-20.45}$
- DFR = $2^{-51.68}$.

Even though the BCH code is used, increasing the message size while maintaining the DFR reduces the security level. However, in Case 6), the security level does not deteriorate that much as compared to Case 4). In addition, it is clear that Gray coding is essential.

Case 7) Frodo-640 with (256, 192, 8) BCH code, Message; 128 →192 bits, DFR $< 2^{-199.56}$:

Encode the 192-bit message with the (256, 192, 8) BCH code to obtain the 256-bit codeword c. c is modulated using Gray coding in (3.2), and then, it passes through the channel. From Table 3.4, B, σ, and the security level are changed as;

- B; $2 \rightarrow 4$
- σ; $2.75 \rightarrow 2.22$
- security level; 149.30 $\rightarrow$ 141.91
- SER = $2^{-26.15}$
- DFR = error probability after BCH decoding =
  probability of more than 8 errors = $2^{-200.70}$.

If the mapping in (3.5) instead of Gray coding in (3.2) is used, SER and DFR are calculated as

- SER = $2^{-26.15}$
- DFR = $2^{-68.79}$.

Even though the BCH code is used, increasing the message size while satisfying DFR $< 2^{-199.56}$ reduces the security level. However, in Case 7), the security level does not reduce much compared to Case 5). In addition, it is clear that Gray coding is essential.

3.2.3 Reducing the Bandwidth of Frodo-640

The bandwidth of Frodo-640 can be reduced by reducing q. Then, σ should be reduced to keep the DFR low because reducing q will increase the DFR. However, there are limits to reducing σ. To meet the bounded distance decoding with the discrete Gaussian sampling (BDDwDGS) reduction requirement, σ should be larger than 2.3 [1].

To reduce the bandwidth of Frodo-640, I can reduce q by half and improve the

security level while satisfying the condition $\sigma \geq 2.3$ using BCH codes, where the DFR still meets the requirement. Table 3.5 summarizes the performances of the following cases for reducing the bandwidth of FrodoKEM schemes.

Table 3.5: Cases for reducing the bandwidth of FrodoKEM scheme.

	q	B	σ	public key [bytes]	ciphertext [bytes]	security level [bits]	DFR with Gray coding	DFR without Gray coding
Frodo-640	32768	2	2.75	9616	9736	149.30	$2^{-148.82}$	·
Case 8)	16384	2	2.3	8976	9088	156.39	$2^{-76.41}$	·
Case 9)	16384	3	2.3	8976	9088	156.39	$2^{-172.63}$	$2^{-59.76}$
Case 10)	8192	3	2.3	8336	8440	172.33	$2^{-28.84}$	$2^{-10.98}$

Case 8) Frodo-640, $q; 32768 \rightarrow 16384, \sigma; 2.75 \rightarrow 2.3$**:**

I reduce q into half without using the BCH code and reduce σ as much as possible to decrease the DFR such as σ = 2.3. Then, SER and DFR are calculated as

- SER is $2^{-82.42}$
- DFR is $2^{-76.41}$.

From these results, it can be seen that the DFR requirement cannot be satisfied by simply reducing q without using the BCH code.

Case 9) Frodo-640 with (192, 128, 8) BCH code, $q; 32768 \rightarrow 16384$**,** $\sigma; 2.75 \rightarrow 2.3$**:**

In this case, $q = 16384$, $\sigma = 2.3$, $B = 3$, and the (192, 128, 8) BCH code are used. The 128-bit message is encoded with the (192, 128, 8) BCH code to obtain the 192-bit codeword c. Then, the codeword c is modulated using Gray coding in (3), and it passes through the channel. Then, SER and DFR are calculated as

- SER = $2^{-23.04}$
- DFR = $2^{-172.63}$.

In this case, the DFR is lower than the DFR requirement $2^{-148.82}$, and the bandwidth of Frodo-640 can be decreased while satisfying the DFR requirement.

- public key; 9616 bytes $\rightarrow$ 8976 bytes
- ciphertext; 9736 bytes $\rightarrow$ 9088 bytes

At this point, the security level is improved as

- security level; 149.30 $\rightarrow$ 156.39.

If the mapping in (2.2) instead of Gray coding in (3.1) is used, SER and DFR are calculated as

- SER = $2^{-23.04}$
- DFR = $2^{-59.76}$.

From these results, it can be seen that the security level can be improved and the bandwidth can be reduced while satisfying the DFR requirement. In addition, it is clear that Gray coding is essential.

Case 10) Frodo-640 with (192, 128, 8) BCH code, $q; 32768 \rightarrow 8192$**,** $\sigma; 2.75 \rightarrow 2.3$**:**

In this case, $q = 8192$, $\sigma = 2.3$, $B = 3$, and the (192, 128, 8) BCH code are used. The 128-bit message is encoded with the (192, 128, 8) BCH code to obtain a 192-bit codeword c. Then, the codeword c is modulated using Gray coding in (3.1), and it passes through the channel. Then, SER and DFR are calculated as

- SER = $2^{-7.00}$
- DFR = $2^{-28.84}$.

In this case, the DFR is higher than the requirement $2^{-148.82}$. If the mapping in

(2.2) instead of Gray coding in (3.1) is used, SER and DFR are calculated as

- SER = $2^{-7.00}$
- DFR = $2^{-10.98}$.

From these results, it can seen that if q is reduced to 8192, then the DFR requirement cannot be satisfied even if the BCH code is used.

Chapter 4

Homomorphic Comparison Using Optimal Composition of Minimax Approximate Polynomials

4.1 Introduction

HE is a cryptographic algorithm that allows algebraic operations over the encrypted data. Until Gentry's seminal work [18] in 2009, HE schemes were able to perform only a few specific operations for the encrypted data. FHE is a cryptographic algorithm that allows all algebraic operations on the encrypted data without restriction and a FHE scheme was first developed in [18]. Due to the feature, FHE has attracted significant attention in various applications and the standardization process for FHE is in progress.

FHE schemes can be classified as bit-wise FHE and word-wise FHE. Word-wise FHE such as Brakerski/Fan-Vercauteren (BFV) [20] and Cheon-Kim-Kim-Song (CKKS) [22] provides the addition and multiplication of an encrypted array over $\mathbb{C}$ or $\mathbb{Z}_p$ for a positive integer $p > 2$. All other operations in word-wise FHE should be performed using these two basic operations. On the other hand, the basic operations of bit-wise FHEs such as TFHE [21] are logic gates such as NAND gates. Recently, word-wise FHE has been widely used in many applications such as deep learning [42, 43].

The comparison function is denoted as $\mathrm{comp}(a, b)$, which outputs 1 if $a > b$, $1/2$ if $a = b$, and 0 if $a < b$. The comparison function is one of the most commonly used operations along with addition and multiplication in many applications including machine learning algorithms [44, 45]. However, when I encrypt inputs word-wise, it is known to be difficult to perform the comparison operation for the ciphertexts in FHEs, called a homomorphic comparison operation, since the comparison operation is a non-polynomial operation. Thus, it is indispensable to find an efficient method to implement the homomorphic comparison operation.

In this dissertation, a new efficient method to perform the homomorphic comparison operation in word-wise FHEs is proposed. Since comparison operation is a non-polynomial operation, it is necessary to find and evaluate a polynomial that approximates $\mathrm{comp}(a, b)$. Comparison operations can be implemented by sign function, that is, $\mathrm{comp}(a, b) = \frac{1}{2}(\mathrm{sgn}(a - b) + 1)$. Thus, in order to perform a homomorphic comparison operation, it is enough to find a polynomial that well approximates $\mathrm{sgn}(x)$.

It is desirable to find the approximate polynomial that requires the minimum computational complexity and depth consumption while satisfying a given approximation error bound. Addition, scalar multiplication, and non-scalar multiplication affect the computational complexity. However, non-scalar multiplication requires the largest computational complexity by far. Although the efficiency of FHE has been improved a lot, non-scalar multiplication still requires a considerable amount of computational complexity. Thus, a polynomial approximation of $\mathrm{sgn}(x)$, which minimizes the number of non-scalar multiplications and depth consumption, is proposed in this dissertation.

4.1.1 Previous Works

Some research has been done on how to find polynomials that approximate the sign function $\mathrm{sgn}(x)$ or $\mathrm{comp}(a, b)$ in FHE. An analytic method to approximate the sign function using the Fourier series was proposed in [46]. In [47], the sign function was

approximated using the approximate equation $\tanh(kx) = \frac{e^{kx}-e^{-kx}}{e^{kx}+e^{-kx}} \simeq \mathrm{sgn}(x)$ for large $k > 0$. Recently, an iterative algorithm was proposed that performs homomorphic comparison operation using the equation $\lim_{k\to\infty} \frac{a^k}{a^k+b^k} = \mathrm{comp}(a, b)$ in [23], where the inverse operation can be performed using Goldschmidt's division algorithm [48]. However, the use of inverse operation causes some inefficiency in computational complexity. More recently, the homomorphic comparison operation is approximated using composite polynomial with less non-scalar multiplications and depth consumption than the previous methods in [24]. It was also shown that the homomorphic comparison operation by using composite polynomial has optimal asymptotic computational complexity. However, the performance of the homomorphic comparison operation using composite polynomials in [24] can be further improved since the composite polynomials used in [24] do not guarantee optimality for the approximation of the sign function by polynomials. Although there have been some improvements, the homomorphic comparison operation still requires a lot of time, and thus more research is needed to improve the performance of the homomorphic comparison operation for practical use.

4.1.2 My Contributions

In this dissertation, I propose that if composite polynomials of component minimax approximate polynomials obtained by the modified Remez algorithm [31] are used, the efficiency of the homomorphic comparison operation can be further improved, where I have three contributions as follows.

First, I propose a method of approximating the sign function with composite polynomials of component minimax approximate polynomials. My main idea is to find the composite polynomial which minimizes the non-scalar multiplications and depth consumption among all of the composite polynomials of component minimax approximate polynomials.

Second, since the sign function is an odd function, it is natural to use composite

polynomials consisting of component polynomials with only odd degree terms. All the component polynomials used in [24] are also polynomials with odd degree terms. It is proved that the composite polynomials of component polynomials with odd degree terms found by the proposed method is the best among all of the composite polynomials of component polynomials with odd degree terms. That is, the composite polynomial obtained by the proposed method requires less number of non-scalar multiplications and depth consumption than any other composite polynomials of component polynomials with odd degree terms.

Third, even though the optimal composite polynomials of component minimax approximate polynomials can be found by the brute-force search from the candidate composite polynomials of component minimax approximate polynomials, the brute-force search requires an exponential time with respect to α, which corresponds to bit precision. Thus, polynomial-time algorithms using dynamic programming which find the optimal composite polynomials in polynomial time are proposed. By using the dynamic programming, the number of required non-scalar multiplications and depth consumption for evaluation of the proposed composite polynomials for the homomorphic comparison operation are obtained and compared to those for the previous method [24]. It can be seen that for the case that I want to minimize the number of non-scalar multiplications, the proposed algorithm reduces the required number of non-scalar multiplications and depth consumption by about 33% and 35%, respectively, compared to those for the previous algorithm. In addition, for the case that I want to minimize the depth consumption, the proposed algorithm reduces the required number of non-scalar multiplications and depth consumption by about 10% and 47%, respectively, compared to those for the previous work.

4.2 Approximation of Sign Function by Using Optimal Composition of Minimax Approximate Polynomials

4.2.1 New Approximation Method for Sine Function Using Composition of the Minimax Approximate Polynomials

In [24], the error of the approximate comparison polynomial compared to the $\mathrm{comp}(a, b)$ is required to be bounded by $2^{-\alpha}$ for any $a, b \in [0, 1]$ satisfying $|a-b| \geq \epsilon$. Note that $\mathrm{comp}(a, b) = \frac{\mathrm{sgn}(a-b)+1}{2}$. If a polynomial $p(x)$ approximating $\mathrm{sgn}(x)$ is $(\alpha - 1, \epsilon)$-close, then the error of $\frac{p(a-b)+1}{2}$ compared to $\mathrm{comp}(a, b)$ is bounded by $2^{-\alpha}$ for any $a, b \in [0, 1]$ satisfying $|a - b| \geq \epsilon$. Thus, I find composite polynomials approximating $\mathrm{sgn}(x)$ that satisfy $(\alpha - 1, \epsilon)$-close to compare the performance of the proposed homomorphic comparison method fairly with that of the previous method in [24].

In [24], $\mathrm{sgn}(x)$ was approximated by using a composite polynomial whose component polynomial is f_n on $[-1, 1]$, which satisfies the following three properties:

(i) $f_n(-x) = -f_n(x)$

(ii) $f_n(1) = 1, f_n(-1) = -1$

(iii) $f_n'(x) = c(1 - x)^n(1 + x)^n$ for some constant $c > 0$.

Then, the only polynomial satisfying the above three properties is given as

$$f_n(x) = \sum_{i=0}^{n} \frac{1}{4^i} \binom{2i}{i} x(1 - x^2)^i.$$

If n and the number of compositions s_n of f_n become larger, the composite polynomial $f_n^{(s_n)}$ approximates $\mathrm{sgn}(x)$ better. In [24], it is stated that they have the best performance when $n = 4$. In addition, by defining and using the other polynomial g_n together with f_n for composition, the efficiency of the composite polynomial is further

improved with the smaller number of the required compositions. However, the polynomial f_n that satisfies the above three properties does not guarantee the optimality for approximation using a composite polynomial. The other polynomial g_n defined in [24] has good properties, but it does not guarantee the optimality, too.

In this dissertation, I construct composite polynomials using new component polynomials f_i's, which are different from those used in the previous paper [24] and the repeated composition of each f_i is not used, that is, $s_i = 1$ for all i. Let $f_k \circ f_{k-1} \circ \cdots \circ f_1$ be a composite polynomial of component polynomials with odd degree terms approximating $\mathrm{sgn}(x)$ on $[-1, -\epsilon] \cup [\epsilon, 1]$. Let $[a_0, b_0] = [\epsilon, 1]$, $f_1([a_0, b_0]) = [a_1, b_1]$, $f_2([a_1, b_1]) = [a_2, b_2], \cdots, f_k([a_{k-1}, b_{k-1}]) = [a_k, b_k]$. Note that $f_k \circ f_{k-1} \circ \cdots \circ f_1$ is $(\alpha - 1, \epsilon)$-close if and only if $f_k \circ f_{k-1} \circ \cdots \circ f_1([\epsilon, 1]) = [a_k, b_k] \subseteq [1 - 2^{1-\alpha}, 1 + 2^{1-\alpha}]$. Since $[a_k, b_k]$ should be a very small interval, it is desirable for each component polynomial f_i on the domain $[a_{i-1}, b_{i-1}]$ to reduce the range as much as possible. My key observation is that if the minimax approximate polynomials are used as component polynomials, the size of the range $[a_i, b_i]$ can be reduced quickly as i increases. Thus, I use a composite polynomial of component minimax approximate polynomials obtained by the modified Remez algorithm.

In this dissertation, the Paterson-Stockmeyer algorithm [49] is used for evaluating the approximate polynomials. Table 4.1 shows the required depth consumption and the number of non-scalar multiplications for evaluating the approximate polynomials with odd degree terms using the Paterson-Stockmeyer algorithm. The exact required number of non-scalar multiplications and the depth consumption differ slightly depending on how the original Paterson-Stockmeyer algorithm [49] is modified. I refer to several papers and find the minimum number of required non-scalar multiplications and the depth consumption among them for each degree. I refer to the values in [24] for polynomials of degree smaller than or equal to 15 and the values in [31] for polynomials of degree larger than or equal to 17. The required depth consumption and the number of non-scalar multiplications for evaluating a polynomial of degree d with odd

degree terms by using the Paterson-Stockmeyer algorithm are denoted by $\mathsf{dep}(d)$ and $\mathsf{mult}(d)$.

The following definitions are necessary for description of Lemma 4.1.

Definition 4.1 ([24])**.** *For $\alpha > 0$ and $0 < \delta < 1$, a polynomial $p(x)$ is said to be (α, δ)-two-sided-close to $\mathrm{sgn}(x)$ if p satisfies the following:*

$$||p(x) - \mathrm{sgn}(x)||_{\infty,[-1-\delta,-1+\delta]\cup[1-\delta,1+\delta]} \leq 2^{-\alpha},$$

where $||\cdot||_{\infty,D}$ denotes the infinity norm over the domain D.

Definition 4.2. *Let $\{f_i\}_{1\leq i\leq k}$ be a set of polynomials satisfying $\deg(f_i) = d_i$, $1 \leq i \leq k$. $\mathsf{MultNum}(\{f_i\}_{1\leq i\leq k})$ and $\mathsf{DepNum}(\{f_i\}_{1\leq i\leq k})$ denote the sum of the numbers of non-scalar multiplications and the sum of depth consumptions required to evaluate f_i for $1 \leq i \leq k$ by using Paterson-Stockmeyer algorithm, respectively. That is,*

$$\mathsf{MultNum}(\{f_i\}_{1\leq i\leq k}) = \sum_{i=1}^{k} \mathsf{mult}(\deg(f_i))$$

$$\mathsf{DepNum}(\{f_i\}_{1\leq i\leq k}) = \sum_{i=1}^{k} \mathsf{dep}(\deg(f_i)).$$

My goal is to find a $(\alpha-1, \epsilon)$-close composite polynomial $f_k \circ f_{k-1} \circ \cdots \circ f_1$ while minimizing $\mathsf{MultNum}(\{f_i\}_{1\leq i\leq k})$ and $\mathsf{DepNum}(\{f_i\}_{1\leq i\leq k})$. The following lemma implies that finding a $(\alpha - 1, \epsilon)$-close composite polynomial $f_k \circ f_{k-1} \circ \cdots \circ f_1$ is equivalent to finding a $(\alpha - 1, \delta)$-two-sided-close composite polynomial $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ when $\delta = \frac{1-\epsilon}{1+\epsilon}$.

Lemma 4.1. *For a set of polynomials with odd degree terms $\{f_i\}_{1\leq i\leq k}$, let $\{\tilde{f}_i\}_{1\leq i\leq k}$ be a set of polynomials with odd degree terms such that $\tilde{f}_1(x) = f_1(\frac{1+\epsilon}{2}x)$ and $\tilde{f}_i(x) = f_i(x)$, $2 \leq i \leq k$. Then, $f_k \circ f_{k-1} \circ \cdots \circ f_1$ is $(\alpha - 1, \epsilon)$-close if and only if $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ is $(\alpha - 1, \delta)$-two-sided-close when $\delta = \frac{1-\epsilon}{1+\epsilon}$.*

Proof. Let $f_k \circ f_{k-1} \circ \cdots \circ f_1$ be a $(\alpha-1, \epsilon)$-close composite polynomial of component polynomials with odd degree terms. Since $f_k \circ f_{k-1} \circ \cdots \circ f_1(x)$ is a polynomial with

Table 4.1: The required depth consumption and the number of non-scalar multiplications for evaluating polynomials with odd degree terms using Paterson-Stockmeyer algorithm [24, 31]

polynomial degree d	$\mathsf{dep}(d)$	$\mathsf{mult}(d)$
3	2	2
5	3	3
7	3	4
9	4	4
11	4	5
13	4	6
15	4	7
17	5	7
19	5	8
21	5	8
23	5	8
25	5	10
27	5	10
29	5	10
31	5	10

odd degree terms, it is sufficient to consider only when $x > 0$. Then, $f_k \circ f_{k-1} \circ \cdots \circ f_1(x) \in [1 - 2^{-(\alpha-1)}, 1 + 2^{-(\alpha-1)}]$ for $\epsilon \leq x \leq 1$. Let $x' = \frac{2}{1+\epsilon}x$. $\epsilon \leq x \leq 1$ corresponds to $1-\delta \leq x' \leq 1+\delta$. Then, $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1(x') = f_k \circ f_{k-1} \circ \cdots \circ f_1(x) \in [1-2^{-(\alpha-1)}, 1+2^{-(\alpha-1)}]$ for $1-\delta \leq x' \leq 1+\delta$. Thus, $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ is $(\alpha-1, \delta)$-two-sided-close. Conversely, let $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1(x') \in [1 - 2^{-(\alpha-1)}, 1 + 2^{-(\alpha-1)}]$ for $1 - \delta \leq x' \leq 1 + \delta$. Let $x = \frac{1+\epsilon}{2}x'$. $1 - \delta \leq x' \leq 1 + \delta$ corresponds to $\epsilon \leq x \leq 1$. Then, $f_k \circ f_{k-1} \circ \cdots \circ f_1(x) = \tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1(x') \in [1 - 2^{-(\alpha-1)}, 1 + 2^{-(\alpha-1)}]$ for $\epsilon \leq x \leq 1$, which means that $f_k \circ f_{k-1} \circ \cdots \circ f_1$ is $(\alpha - 1, \epsilon)$-close. Thus, the lemma is proved. □ □

Note that since $\deg(f_i) = \deg(\tilde{f}_i), 1 \leq i \leq k$ in Lemma 4.1, it holds that

$$\mathsf{MultNum}(\{f_i\}_{1\leq i\leq k}) = \mathsf{MultNum}(\{\tilde{f}_i\}_{1\leq i\leq k})$$
$$\mathsf{DepNum}(\{f_i\}_{1\leq i\leq k}) = \mathsf{DepNum}(\{\tilde{f}_i\}_{1\leq i\leq k}).$$

Thus, for any $m, n \in \mathbb{N}$, a composite polynomial of component polynomials with odd degree terms $f_k \circ f_{k-1} \circ \cdots \circ f_1$ is $(\alpha-1, \epsilon)$-close and satisfies $\mathsf{MultNum}(\{f_i\}_{1\leq i\leq k}) = m$ and $\mathsf{DepNum}(\{f_i\}_{1\leq i\leq k}) = n$ if and only if the corresponding composite polynomial $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ is $(\alpha - 1, \delta)$-two-sided-close and satisfies $\mathsf{MultNum}(\{\tilde{f}_i\}_{1\leq i\leq k}) = m$ and $\mathsf{DepNum}(\{\tilde{f}_i\}_{1\leq i\leq k}) = n$ when $\delta = \frac{1-\epsilon}{1+\epsilon}$. Thus, it can be seen that the following two algorithms are equivalent:

(i) An algorithm that finds the $(\alpha - 1, \epsilon)$-close composite polynomial $f_k \circ \cdots \circ f_1$ which minimizes the number of non-scalar multiplications and the depth consumption

(ii) An algorithm that finds the $(\alpha - 1, \delta)$-two-sided-close composite polynomial $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ which minimizes the number of non-scalar multiplications and the depth consumption

Thus, from now on, I focus on finding $(\alpha - 1, \delta)$-two-sided-close composite polynomial $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ which minimizes the number of non-scalar multiplications and the depth consumption.

The minimax composite polynomial, which is the core of the proposed homomorphic comparison method, is now defined as follows. The main idea of the proposed approximation method is to use the minimax composite polynomial to approximate the sign function. I denote $[-1-s, -1+s] \cup [1-s, 1+s]$ by R_s for $s > 0$.

Definition 4.3. *Let $\{f_i\}_{1 \le i \le k}$ be a set of polynomials. Let D be $[-b, -a] \cup [a, b]$. $f_k \circ f_{k-1} \circ \cdots \circ f_1$ is called a minimax composite polynomial on D if there exists $\{d_i\}_{1 \le i \le k}$ that satisfies the followings:*

- *f_1 is the minimax approximate polynomial of degree at most d_1 on D for $\mathrm{sgn}(x)$ and the minimax approximation error is equal to τ_1.*
- *For $2 \le i \le k$, f_i is the minimax approximate polynomial of degree at most d_i on $f_{i-1} \circ f_{i-2} \circ \cdots \circ f_1(D)$ for $\mathrm{sgn}(x)$. The minimax approximation error is τ_i.*

Note that $f_i \circ f_{i-1} \circ \cdots \circ f_1(D) = R_{\tau_i}$, $1 \le i \le k$ from Theorem 2.1. In fact, τ_i becomes smaller as i increases. It can be seen that if $f_k \circ f_{k-1} \circ \cdots \circ f_1$ is a minimax composite polynomial on $D = [-b, -a] \cup [a, b]$, then $\{f_i\}_{1 \le i \le k}$ is a set of polynomials with odd degree terms from Lemma 2.2. If $\tau_k \le 2^{-(\alpha-1)}$, then the minimax composite polynomial on R_δ becomes $(\alpha - 1, \delta)$-two-sided-close. My key idea is to find the minimax composite polynomial on R_δ that requires the minimum number of non-scalar multiplications and depth consumption among all $(\alpha - 1, \delta)$-two-sided-close minimax composite polynomials on R_δ. Note that there is a tradeoff between the number of non-scalar multiplications and the depth consumption. I deal with both cases when putting priority on minimizing the number of non-scalar multiplications and on minimizing the depth consumption.

4.2.2 Optimality of Approximation of the Sign Function by a Minimax Composite Polynomial

Since $\mathrm{sgn}(x)$ is an odd function, it is natural to approximate $\mathrm{sgn}(x)$ by using a composite polynomial of component polynomials with odd degree terms. Assume that I can obtain the minimax composite polynomial on R_δ that requires the minimum number of non-scalar multiplications and depth consumption among all minimax composite polynomials on R_δ satisfying $(\alpha - 1, \delta)$-two-sided-close. In this subsection, it is proved that the obtained minimax composite polynomial on R_δ requires less number of non-scalar multiplications and depth consumption than any $(\alpha - 1, \delta)$-two-sided-close composite polynomial of component polynomials with odd degree terms. That is, for any $(\alpha - 1, \delta)$-two-sided-close composite polynomial of component polynomials with odd degree terms, there exists a $(\alpha - 1, \delta)$-two-sided-close minimax composite polynomial on R_δ such that the number of required non-scalar multiplications and the depth consumption for the minimax composite polynomial are less than or equal to those for the composite polynomial of component polynomials with odd degree terms, respectively.

The following definition and lemmas are needed for the proof of optimality of the proposed approximation method of approximating the sign function using a minimax composite polynomial.

Definition 4.4. *Let $\{f_i\}_{1\leq i\leq k}$ be a set of polynomials. $f_k \circ f_{k-1} \circ \cdots \circ f_1$ is called a 1-centered range composite polynomial on R_δ if $\{f_i\}_{1\leq i\leq k}$ is a set of polynomials with odd degree terms and there exists $\{\tau_i\}_{1\leq i\leq k}$ such that $f_1([1-\delta, 1+\delta]) = [1-\tau_1, 1+\tau_1]$ and $f_i([1-\tau_{i-1}, 1+\tau_{i-1}]) = [1-\tau_i, 1+\tau_i]$ for $2 \leq i \leq k$.*

Lemma 4.2. *Let f_1 be the minimax approximate polynomial of degree at most d on $[-b_1, -a_1] \cup [a_1, b_1]$ for $\mathrm{sgn}(x)$. Let f_2 be the minimax approximate polynomial of degree at most d on $[-b_2, -a_2] \cup [a_2, b_2]$ for $\mathrm{sgn}(x)$. If $[a_2, b_2] \subseteq [a_1, b_1]$, then the minimax approximation error e_2 of f_2 is less than or equal to the minimax approxima-*

tion error e_1 of f_1.

Proof. When f_1 approximates $\mathrm{sgn}(x)$ on $[-b_1, -a_1] \cup [a_1, b_1]$, the maximum approximation error e_1 is larger than or equal to the maximum approximation error e_1' when f_1 approximates $\mathrm{sgn}(x)$ on $[-b_2, -a_2] \cup [a_2, b_2]$. According to the definition of minimax approximate polynomial, f_2 is the polynomial with the smallest maximum approximation error when approximating $\mathrm{sgn}(x)$ on $[-b_2, -a_2] \cup [a_2, b_2]$ among all polynomials of degree smaller than or equal to d. Among polynomials of degree smaller than or equal to d, there is also f_1. Thus, it holds that $e_2 \leq e_1' \leq e_1$, and the lemma is proved. □ □

Lemma 4.3. *Let $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ be any $(\alpha-1, \delta)$-two-sided-close 1-centered range composite polynomial on R_δ. Then, there is a $(\alpha - 1, \delta)$-two-sided-close minimax composite polynomial $\hat{f}_k \circ \hat{f}_{k-1} \circ \cdots \circ \hat{f}_1$ on R_δ such that $\deg(\hat{f}_i) \leq \deg(\tilde{f}_i)$ for i, $1 \leq i \leq k$.*

Proof. Since $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ is a 1-centered range composite polynomial on R_δ, there exists $\{\tau_i\}_{1 \leq i \leq k}$ such that $\tilde{f}_1([1-\delta, 1+\delta]) = [1-\tau_1, 1+\tau_1]$ and $\tilde{f}_i([1-\tau_{i-1}, 1+\tau_{i-1}) = [1-\tau_i, 1+\tau_i]$ for all i, $2 \leq i \leq k$. Then, $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1([1-\delta, 1+\delta]) = [1-\tau_k, 1+\tau_k]$. Since $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ is $(\alpha-1, \delta)$-two-sided-close, $\tau_k \leq 2^{-(\alpha-1)}$ should hold. Let $\deg(\tilde{f}_i) = d_i$, $1 \leq i \leq k$. Let $\hat{f}_1$ be the minimax approximate polynomial of degree at most d_1 on R_δ and let τ_1' be the approximation error of $\hat{f}_1$. Then $\tau_1' \leq \tau_1$. Let τ_i' be the approximation error of $\hat{f}_i$, which is the minimax approximate polynomial of degree at most d_i on $R_{\tau_{i-1}'}$ for $\mathrm{sgn}(x)$ for i, $2 \leq i \leq k$. Then, $\hat{f}_k \circ \hat{f}_{k-1} \circ \cdots \circ \hat{f}_1$ is a minimax composite polynomial on R_δ. I want to show that $\tau_i' \leq \tau_i$, $2 \leq i \leq k$ by inductive method. Assume that $\tau_{i-1}' \leq \tau_{i-1}$. Let τ_i'' be the approximation error of the minimax approximate polynomial of degree at most d_i on $R_{\tau_{i-1}}$ for $\mathrm{sgn}(x)$. From Lemma 4.2, it holds that $\tau_i' \leq \tau_i''$. Since $\tilde{f}_i([1-\tau_{i-1}, 1+\tau_{i-1}]) = [1-\tau_i, 1+\tau_i]$, $\tau_i'' \leq \tau_i$ holds. Thus, $\tau_i' \leq \tau_i'' \leq \tau_i$. It holds that $\tau_i' \leq \tau_i$ for all i, $2 \leq i \leq k$ by inductive method. Since $\tau_k' \leq \tau_k \leq 2^{-(\alpha-1)}$, $\hat{f}_k \circ \hat{f}_{k-1} \circ \cdots \circ \hat{f}_1$ is a $(\alpha-1, \delta)$-two-

sided-close minimax composite polynomial on R_δ such that $\deg(\hat{f}_i) \leq \deg(\tilde{f}_i)$ for all $i, 1 \leq i \leq k$. □ □

Lemma 4.4. *Let $f_k \circ f_{k-1} \circ \cdots \circ f_1$ be any $(\alpha - 1, \delta)$-two-sided-close composite polynomial of component polynomials with odd degree terms. Then, there is a $(\alpha - 1, \delta)$-two-sided-close 1-centered range composite polynomial $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ on R_δ such that $\deg(\tilde{f}_i) = \deg(f_i)$ for all i, $1 \leq i \leq k$.*

Proof. Let $f_1([1-\delta, 1+\delta]) = [a_1, b_1], f_2([a_1, b_1]) = [a_2, b_2], \cdots, f_k([a_{k-1}, b_{k-1}]) = [a_k, b_k]$. Since $\{f_i\}_{1 \leq i \leq k}$ is a set of polynomials with odd degree terms, it holds that $f_1([-1 - \delta, -1 + \delta]) = [-b_1, -a_1], f_2([-b_1, -a_1]) = [-b_2, -a_2], \cdots, f_k([-b_{k-1}, -a_{k-1}]) = [-b_k, -a_k]$. Satisfying $(\alpha - 1, \delta)$-two-sided-close means that $[a_k, b_k] \subseteq [1 - 2^{-(\alpha-1)}, 1 + 2^{-(\alpha-1)}]$. Also, it is easy to see that $0 < a_i < b_i$ for $i, 1 \leq i \leq k$ from the fact that $f_k \circ f_{k-1} \circ \cdots \circ f_1$ is a $(\alpha - 1, \delta)$-two-sided-close composite polynomial. Let $\tilde{f}_1(x) = \frac{2}{a_1+b_1} f_1(x)$ and $\tilde{f}_i(x) = \frac{2}{a_i+b_i} f_i(\frac{a_i+b_i}{2}x)$, $2 \leq i \leq k$. Then, $\tilde{f}_1([1-\delta, 1+\delta]) = [1 - \frac{b_1-a_1}{a_1+b_1}, 1 + \frac{b_1-a_1}{a_1+b_1}]$ and $\tilde{f}_i([1 - \frac{b_{i-1}-a_{i-1}}{a_{i-1}+b_{i-1}}, 1 + \frac{b_{i-1}-a_{i-1}}{a_{i-1}+b_{i-1}}]) = [1 - \frac{b_i-a_i}{a_i+b_i}, 1 + \frac{b_i-a_i}{a_i+b_i}]$, $2 \leq i \leq k$. Then, $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ is a 1-centered range composite polynomial on R_δ. I want to show that $[1 - \frac{b_k-a_k}{a_k+b_k}, 1 + \frac{b_k-a_k}{a_k+b_k}] = [\frac{2a_k}{a_k+b_k}, \frac{2b_k}{a_k+b_k}] \subseteq [1 - 2^{-(\alpha-1)}, 1 + 2^{-(\alpha-1)}]$, which means that $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ is $(\alpha - 1, \delta)$-two-sided-close. If $a_k + b_k \leq 2$, then $1 - 2^{-(\alpha-1)} \leq a_k \leq \frac{2a_k}{a_k+b_k}$ and $\frac{2b_k}{a_k+b_k} = 2 - \frac{2a_k}{a_k+b_k} \leq 2 - a_k \leq 1 + 2^{-(\alpha-1)}$ hold. On the other hand, if $a_k + b_k > 2$, then $\frac{2b_k}{a_k+b_k} < b_k \leq 1 + 2^{-(\alpha-1)}$ and $\frac{2a_k}{a_k+b_k} = 2 - \frac{2b_k}{a_k+b_k} > 2 - b_k \geq 1 - 2^{-(\alpha-1)}$ hold. Thus $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ is a $(\alpha - 1, \delta)$-two-sided-close 1-centered range composite polynomial on R_δ satisfying $\deg(\tilde{f}_i) = \deg(f_i)$ for all i, $1 \leq i \leq k$. □ □

The following procedure for proof of Theorem 4.1 is used:

(i) It is proved that for any $(\alpha - 1, \delta)$-two-sided-close 1-centered range composite polynomial $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ on R_δ, it holds that $\deg(\hat{f}_i) = \deg(\tilde{f}_i)$ for all i,

$1 \leq i \leq k$ for some $(\alpha - 1, \delta)$-two-sided-close minimax composite polynomial $\hat{f}_k \circ \hat{f}_{k-1} \circ \cdots \circ \hat{f}_1$ on R_δ from Lemma 4.3.

(ii) It is proved that for any $(\alpha - 1, \delta)$-two-sided-close composite polynomial of component polynomials with odd degree terms $\{f_i\}_{1 \leq i \leq k}$, it holds that $\deg(\tilde{f}_i) \leq \deg(f_i)$ for all i, $1 \leq i \leq k$ for some $(\alpha - 1, \delta)$-two-sided-close 1-centered range composite polynomial $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ on R_δ from Lemma 4.4.

(iii) Finally, with above lemmas, it is proved in Theorem 4.1 that for any $(\alpha - 1, \delta)$-two-sided-close composite polynomial of component polynomials with odd degree terms $\{f_i\}_{1 \leq i \leq k}$, it holds that $\deg(\hat{f}_i) \leq \deg(f_i)$ for all i, $1 \leq i \leq k$ for some $(\alpha - 1, \delta)$-two-sided-close minimax composite polynomial $\hat{f}_k \circ \hat{f}_{k-1} \circ \cdots \circ \hat{f}_1$ on R_δ.

Theorem 4.1. *Let $f_k \circ f_{k-1} \circ \cdots \circ f_1$ be any $(\alpha - 1, \delta)$-two-sided-close composite polynomial of component polynomials with odd degree terms. Then, there is a $(\alpha - 1, \delta)$-two-sided-close minimax composite polynomial $\hat{f}_k \circ \hat{f}_{k-1} \circ \cdots \circ \hat{f}_1$ on R_δ such that $\deg(\hat{f}_i) \leq \deg(f_i)$ for all i, $1 \leq i \leq k$.*

Proof. Let $f_k \circ f_{k-1} \circ \cdots \circ f_1$ be any $(\alpha - 1, \delta)$-two-sided-close composite polynomial of component polynomials with odd degree terms. From Lemma 4.4, there is a $(\alpha - 1, \delta)$-two-sided-close 1-centered range composite polynomial $\tilde{f}_k \circ \tilde{f}_{k-1} \circ \cdots \circ \tilde{f}_1$ on R_δ such that $\deg(\tilde{f}_i) = \deg(f_i)$, $1 \leq i \leq k$. In addition, from Lemma 4.3, there is a $(\alpha - 1, \delta)$-two-sided-close minimax composite polynomial $\hat{f}_k \circ \hat{f}_{k-1} \circ \cdots \circ \hat{f}_1$ on R_δ such that $\deg(\hat{f}_i) \leq \deg(\tilde{f}_i)$, $1 \leq i \leq k$. Thus, there is a $(\alpha - 1, \delta)$-two-sided-close minimax composite polynomial $\hat{f}_k \circ \hat{f}_{k-1} \circ \cdots \circ \hat{f}_1$ on R_δ such that $\deg(\hat{f}_i) \leq \deg(f_i)$ for all i, $1 \leq i \leq k$. □ □

Remark 2. *In Theorem 4.1, since $\deg(\hat{f}_i) \leq \deg(f_i)$ for $1 \leq i \leq k$, it holds that* $\mathsf{MultNum}(\{\hat{f}_i\}_{1 \leq i \leq k}) \leq \mathsf{MultNum}(\{f_i\}_{1 \leq i \leq k})$ *and* $\mathsf{DepNum}(\{\hat{f}_i\}_{1 \leq i \leq k}) \leq$

$\mathsf{DepNum}(\{f_i\}_{1\leq i\leq k})$. *It means that if I find the minimax composite polynomial on* R_δ *that requires the minimum number of non-scalar multiplications and depth consumption among all* $(\alpha-1,\delta)$*-two-sided-close minimax composite polynomials on* R_δ*, the number of required non-scalar multiplications and depth consumption for the obtained minimax composite polynomial on* R_δ *are less than or equal to those for any* $(\alpha-1,\delta)$*-two-sided-close composite polynomial of component polynomials with odd degree terms, respectively.*

4.2.3 Achieving Polynomial-Time Algorithm for New Approximation Method by Using Dynamic Programming

I can find the $(\alpha-1,\delta)$-two-sided-close minimax composite polynomial on R_δ that requires the minimum number of non-scalar multiplications and depth consumption among all $(\alpha-1,\delta)$-two-sided-close minimax composite polynomials by brute-force search. However, the brute-force search requires considerable time. Thus, dynamic programming is used to find the minimax composite polynomial on R_δ with the computational complexity in polynomial time. Thus, I propose an algorithm to find the minimax composite polynomial on R_δ that requires the minimum number of non-scalar multiplications and depth consumption in polynomial time by using dynamic programming.

$\mathsf{MinErr}(d,t)$, $\mathsf{InvMinErr}(d,t)$, $f(m,n,t)$, and $G(m,n,t)$ are defined before the description of the proposed algorithms as follows.

Definition 4.5. *For* $d\in\mathbb{N}$ *and* $t\in(0,1)$, $\mathsf{MinErr}(d,t)$ *is the minimax approximation error of the minimax approximate polynomial of degree at most* d *on* R_t *for* $\mathrm{sgn}(x)$.

Lemma 4.5. *For a fixed odd* $d\in\mathbb{N}$, $\mathsf{MinErr}(d,t)$ *is a strictly increasing continuous function of* t *on* $(0,1)$.

Proof. Let d be $2i+1$. Consider the minimax approximate polynomial $p(x)$ of degree at most $2i+1$ on R_t for $\mathrm{sgn}(x)$. Let τ_0 be the minimax approximaton error of $p(x)$

on R_t. Since $\mathrm{sgn}(x)$ is an odd function, it can be seen from Lemma 2.2 that the minimax approximate polynomial of degree at most $2i+1$ to $\mathrm{sgn}(x)$ is equal to the minimax approximate polynomial of degree at most $2i+2$ to $\mathrm{sgn}(x)$. Also, $p(x)$ is a polynomial with odd degree terms from Lemma 2.2. I want to show that there exist $i+2$ distinct points $x_0, x_1, \cdots, x_{i+1} \in [1-t, 1+t]$ that satisfy the following three properties:

Prop 1. $1-t = x_0 < x_1 < \cdots < x_{i+1} = 1+t$.

Prop 2. $p(x_j) = 1 + (-1)^{j+1}\tau_0$, $0 \le j \le i+1$.

Prop 3. $p(x)$ is strictly increasing on $(0, x_1)$. For j, $1 \le j \le i$, $p(x)$ is strictly increasing on (x_j, x_{j+1}) when j is even, and strictly decreasing on (x_j, x_{j+1}) when j is odd. Also, $p(x)$ is strictly increasing on (x_i, ∞) if i is even and strictly decreasing on (x_i, ∞) if i is odd.

$|p(x) - \mathrm{sgn}(x)|$ should have maximum values at $2i+4$ distinct points in R_t from Theorem 2.1. However, there are at most $2i$ distinct points x such that $p'(x) = 0$. If I consider when $x > 0$, $|p(x) - \mathrm{sgn}(x)|$ should have maximum values at $i+2$ distinct points on $[1-t, 1+t]$ and there are at most i distinct points x such that $p'(x) = 0$. If $|p(x) - \mathrm{sgn}(x)|$ has maximum value at $x = x_0$, then it holds that $p'(x_0) = 0$ or $x = x_0$ is a boundary point, that is, $x_0 \in \{1-t, 1+t\}$. Thus, $|p(x) - \mathrm{sgn}(x)|$ should have maximum values at two boundary points $x = 1-t$ and $x = 1+t$. Let $x_0, \cdots, x_{i+1}$ be the $i+2$ distinct points on $(0, \infty)$ such that $|p(x) - \mathrm{sgn}(x)|$ has maximum values at those points. Then, it holds that $x_0 = 1-t, x_{i+1} = 1+t$ and $p'(x_1) = p'(x_2) = \cdots, p'(x_i) = 0$. Also, considering $p(0) = 0$ and $p(x_1) > 0$, $p(x)$ is strictly increasing on $(0, x_1)$. Since $p(x_0) < p(x_1)$, it holds that $p(x_0) = 1 - \tau_0, p(x_1) = 1 + \tau_0, p(x_2) = 1 - \tau_0, \cdots$ from Theorem 2.1. Also, it can be seen that the Prop 3 is satisfied from Theorem 2.1. Thus, there exist $i+2$ points $x_0, x_1, \cdots, x_{i+1} \in (0, \infty)$ that satisfy the above three properties. Now I want to show

that $\mathsf{MinErr}(d,t)$ is a strictly increasing continuous function of t with domain $(0,1)$ as follows:

(i) Strictly increasing:

Let $0 < t_1 < t_2 < 1$. Let $p_1(x)$ and $p_2(x)$ be the minimax approximate polynomials of degree at most $2i+1$ on R_{t_1} and R_{t_2}, respectively. It is trivial that $\mathsf{MinErr}(d,t_1) \leq \mathsf{MinErr}(d,t_2)$. Assume that $\mathsf{MinErr}(d,t_1) = \mathsf{MinErr}(d,t_2) = \tau_0$. Then, by the uniqueness property of the minimax approximate polynomial, it should be hold that $p_1(x) = p_2(x)$. Note that $p_1(x)$ is the minimax approximate polynomial of degree at most $2i+1$ on R_{t_1}. Then, it can be seen that $0 < p_1(1-t_2) < p_1(1-t_1) = 1-\tau_0$ from Prop 3. Considering $p_1(x) = p_2(x)$, the minimax approximation error of $p_2(x)$ on R_{t_2} is larger than τ_0. That is, $\mathsf{MinErr}(d,t_1) < \mathsf{MinErr}(d,t_2)$, which is a contradiction. Thus, $\mathsf{MinErr}(d,t)$ is a strictly increasing function of t.

(ii) Continuous:

I want to show that $\mathsf{MinErr}(d,t)$ is continuous at $t = t_0$, that is, for any $\delta' > 0$, there exists $\epsilon' > 0$ such that $|t-t_0| \leq \epsilon'$ implies $|\mathsf{MinErr}(d,t) - \mathsf{MinErr}(d,t_0)| \leq \delta'$. Let $p(x)$ be the minimax approximate polynomial of degree at most $2i+1$ on R_{t_0}, and let τ_0 be the minimax approximation error of $p(x)$. It is enough to consider only the case when $\delta' < \tau_0$. There exist $i+2$ distinct points $x_0, x_1, \cdots, x_{i+1} \in (0,\infty)$ that satisfy the above three properties. There exists a unique $x \in (0, x_0)$ such that $p(x) = 1-\tau_0-\delta'$. Let ϵ_1' be $1-t_0-x$ for the unique x. Also, there exists unique $x \in (x_{i+1}, \infty)$ such that $p(x) = 1+\tau_0+\delta'$ when i is even and unique $x \in (x_{i+1}, \infty)$ such that $p(x) = 1-\tau_0-\delta'$ when i is odd. Let ϵ_2' be $x-1-t_0$ for the unique x. There exists unique $x \in (x_0, x_1)$ such that $p(x) = 1-\tau_0+\delta'$. Let ϵ_3' be $x-1+t_0$ for the unique x. Also, there exists unique $x \in (x_i, x_{i+1})$ such that $p(x) = 1+\tau_0-\delta'$ when i is even and unique $x \in (x_i, x_{i+1})$ such that $p(x) = 1-\tau_0+\delta'$. Let

ϵ'_4 be $-x + 1 + t_0$ for the unique x. Now, let ϵ' be $\min(\epsilon'_1, \epsilon'_2, \epsilon'_3, \epsilon'_4)$. Then, $p([1 - t_0 - \epsilon', 1 + t_0 + \epsilon']) \subseteq [1 - \tau_0 - \delta', 1 + \tau_0 + \delta']$. Thus, the minimax approximation error of the minimax approximate polynomial on $R_{t_0+\epsilon'}$ is smaller than or equal to $\tau_0 + \delta'$. That is, $\mathsf{MinErr}(d, t_0 + \epsilon') \leq \tau_0 + \delta'$. On the other hand, let $x'_0 = x_0 + \epsilon', x'_1 = x_1, \cdots, x'_i = x_i, x'_{i+1} = x_{i+1} - \epsilon'$. Consider $2i + 4$ points $-x'_{i+1}, -x'_i, \cdots, -x'_0, x'_0, \cdots, x'_i, x'_{i+1}$. From Lemma 2.1, the minimax approximation error of the minimax approximate polynomial on $R_{t_0-\epsilon'}$ is larger than or equal to $\tau_0 - \delta'$. That is, $\mathsf{MinErr}(d, t_0 - \epsilon') \geq \tau_0 - \delta'$. Since $\mathsf{MinErr}(d, t)$ is an increasing function, if $|t - t_0| \leq \epsilon'$, then $|\mathsf{MinErr}(d, t) - \mathsf{MinErr}(d, t_0)| \leq \delta'$. Thus, $\mathsf{MinErr}(d, t)$ is continuous at $t = t_0$.

□ □

If the minimax approximate polynomial of degree at most d on R_t narrows the domain R_t to a range R_τ, $\mathsf{MinErr}(d, t)$ outputs τ. Since $\mathsf{MinErr}(d, t)$ is strictly increasing function of t on $(0, \infty)$, the inverse function of $\mathsf{MinErr}(d, t)$ exists, which is defined as follows.

Definition 4.6. *For $d \in \mathbb{N}$, $\mathsf{InvMinErr}(d, t)$ is $\tau > 0$ such that $\mathsf{MinErr}(d, \tau) = t$.*

The approximate value of $\mathsf{InvMinErr}(d, t)$ can be obtained by binary search using modified Remez algorithm.

Definition 4.7. *$f(m, n, t)$ is the maximum $\tau \in (0, 1)$ such that there exists a minimax composite polynomial $f_k \circ f_{k-1} \circ \cdots \circ f_1$ on R_τ satisfying $f_k \circ f_{k-1} \circ \cdots \circ f_1([1-\tau, 1+\tau]) \subseteq [1-t, 1+t]$, $\mathsf{MultNum}(\{f_i\}_{1 \leq i \leq k}) \leq m$, and $\mathsf{DepNum}(\{f_i\}_{1 \leq i \leq k}) \leq n$.*

$f(m, n, t)$ outputs the maximum $\tau > 0$ when the range of a minimax composite polynomial on R_τ becomes smaller than R_t with m or less number of non-scalar multiplications and with n or less depth consumption. The degrees of k component polynomials for the corresponding minimax composite polynomial $f_k \circ f_{k-1} \circ \cdots \circ f_1$ on R_τ in Definition 4.7 are stored in $G(m, n, t)$ as an ordered set. It is trivial that if

$0 \le m \le 1$ or $0 \le n \le 1$, then $f(m,n,t) = t$. For $m \ge 2$ and $n \ge 2$, the following theorem for $f(m,n,t)$ holds:

Theorem 4.2. *For $m \ge 2$ and $n \ge 2$, the following recursion for $f(m,n,t)$ holds:*

$$f(m,n,t) = \max_{\substack{1 \le k \\ \mathsf{mult}(2k+1) \le m \\ \mathsf{dep}(2k+1) \le n}} \mathsf{InvMinErr}(2k+1, f(m - \mathsf{mult}(2k+1), n - \mathsf{dep}(2k+1), t)).$$

Proof. Let $\tau = f(m,n,t)$. Assume that

$$\tau > \max_{\substack{1 \le k \\ \mathsf{mult}(2k+1) \le m \\ \mathsf{dep}(2k+1) \le n}} \mathsf{InvMinErr}(2k+1, f(m - \mathsf{mult}(2k+1), n - \mathsf{dep}(2k+1), t)).$$

Then there exists a minimax composite polynomial $f_k \circ f_{k-1} \circ \cdots \circ f_1$ satisfying $f_k \circ f_{k-1} \circ \cdots \circ f_1([1-\tau, 1+\tau]) \subseteq [1-t, 1+t], \mathsf{MultNum}(\{f_i\}_{1 \le i \le k}) \le m$, and $\mathsf{DepNum}(\{f_i\}_{1 \le i \le k}) \le n$. Let d_1 be the degree of f_1 and let $f_1([1-\tau, 1+\tau]) = [1-\tau', 1+\tau']$. Since the minimax composite polynomial $f_k \circ f_{k-1} \circ \cdots \circ f_2$ on $[1-\tau', 1+\tau']$ satisfies $f_k \circ f_{k-1} \circ \cdots \circ f_2([1-\tau', 1+\tau']) \subseteq [1-t, 1+t]$, $\mathsf{MultNum}(\{f_i\}_{2 \le i \le k}) \le m - \mathsf{mult}(d_1)$, and $\mathsf{DepNum}(\{f_i\}_{2 \le i \le k}) \le n - \mathsf{dep}(d_1)$, it holds that $\tau' \le f(m - \mathsf{mult}(d_1), n - \mathsf{dep}(d_1), t)$. Then,

$$\tau = \mathsf{InvMinErr}(d_1, \tau') \le \mathsf{InvMinErr}(d_1, f(m - \mathsf{mult}(d_1), n - \mathsf{dep}(d_1), t))$$
$$\le \max_{\substack{1 \le k \\ \mathsf{mult}(2k+1) \le m \\ \mathsf{dep}(2k+1) \le n}} \mathsf{InvMinErr}(2k+1, f(m - \mathsf{mult}(2k+1), n - \mathsf{dep}(2k+1), t))$$

This leads to a contradiction because

$$\tau > \max_{\substack{1 \le k \\ \mathsf{mult}(2k+1) \le m \\ \mathsf{dep}(2k+1) \le n}} \mathsf{InvMinErr}(2k+1, f(m - \mathsf{mult}(2k+1), n - \mathsf{dep}(2k+1), t)).$$

Assume that

$$\tau < \max_{\substack{1\le k \\ \mathsf{mult}(2k+1)\le m \\ \mathsf{dep}(2k+1)\le n}} \mathsf{InvMinErr}(2k+1, f(m-\mathsf{mult}(2k+1), n-\mathsf{dep}(2k+1), t)).$$

$$\max_{\substack{1\le k \\ \mathsf{mult}(2k+1)\le m \\ \mathsf{dep}(2k+1)\le n}} \mathsf{InvMinErr}(2k+1, f(m-\mathsf{mult}(2k+1), n-\mathsf{dep}(2k+1), t)) =$$

$\mathsf{InvMinErr}(2i+1, f(m-\mathsf{mult}(2i+1), n-\mathsf{dep}(2i+1), t))$ for some i. Let τ' be $\mathsf{InvMinErr}(2i+1, f(m-\mathsf{mult}(2i+1), n-\mathsf{dep}(2i+1), t))$. Let τ'' be $f(m-\mathsf{mult}(2i+1), n-\mathsf{dep}(2i+1), t) = \mathsf{MinErr}(2i+1, \tau')$. Then, there exists a minimax composite polynomial $f_k \circ f_{k-1} \circ \cdots \circ f_2$ satisfying

$$f_k \circ f_{k-1} \circ \cdots \circ f_2([1-\tau'', 1+\tau'']) \subseteq [1-t, 1+t]$$
$$\mathsf{MultNum}(\{f_i\}_{2\le i\le k}) \le m - \mathsf{mult}(2i+1)$$
$$\mathsf{DepNum}(\{f_i\}_{2\le i\le k}) \le n - \mathsf{dep}(2i+1).$$

Let f_1 be the minimax approximate polynomial of degree at most $2i+1$ on $[1-\mathsf{InvMinErr}(2i+1, \tau''), 1+\mathsf{InvMinErr}(2i+1, \tau'')]$. Since $f_1([1-\mathsf{InvMinErr}(2i+1, \tau''), 1+\mathsf{InvMinErr}(2i+1, \tau'')]) = [1-\tau'', 1+\tau'']$, it holds that $f_k \circ f_{k-1} \circ \cdots \circ f_1([1-\mathsf{InvMinErr}(2i+1, \tau''), 1+\mathsf{InvMinErr}(2i+1, \tau'')]) \subseteq [1-t, 1+t]$. Also, it holds that $\mathsf{MultNum}(f_k \circ f_{k-1} \circ \cdots \circ f_1) \le m$ and $\mathsf{DepNum}(f_k \circ f_{k-1} \circ \cdots \circ f_1) \le n$. Thus, $\tau = f(m,n,t) < \mathsf{InvMinErr}(2i+1, \tau'')$, which is a contradiction. Thus, it holds that

$$\tau = \max_{\substack{1\le k \\ \mathsf{mult}(2k+1)\le m \\ \mathsf{dep}(2k+1)\le n}} \mathsf{InvMinErr}(2k+1, f(m-\mathsf{mult}(2k+1), n-\mathsf{dep}(2k+1), t)) \quad (4.1)$$

and the theorem is proved. □ □

$f(m,n,t)$ and $G(m,n,t)$ are recursively computed by the following Algorithm 6. In the 9th line of Algorithm 6, $\{2j+1\} \cup G(m-\mathsf{mult}(2j+1), n-\mathsf{dep}(2j+1), t)$

means inserting $2j+1$ to the ordered set $G(m - \mathsf{mult}(2j+1), n - \mathsf{dep}(2j+1), t)$ as the first component. In this dissertation, only minimax approximate polynomials of degree at most 31 are used to reduce the time complexity of the proposed algorithms. Since numerical results show that only minimax approximate polynomials of degree at most 11 are used to minimize the number of non-scalar multiplications, it seems that the minimax approximate polynomials of degree at most 31 are sufficient when minimizing the number of non-scalar multiplications. On the other hand, minimax approximate polynomials of large degree are sometimes used when minimizing the depth consumption. Thus, if minimax approximate polynomials of degree larger than 31 are also used, the required depth consumption may be further reduced.

Now, **`DynMinMult`** and **`DynMinDep`** algorithms are introduced, which use the values of $f(m,n,t)$ and $G(m,n,t)$ obtained from Algorithm 6. The following two cases are considered, which correspond to **`DynMinMult`** and **`DynMinDep`**, respectively.

First, **`DynMinMult`** puts more priority on minimizing the number of non-scalar multiplications rather than minimizing the depth consumption. The minimum number of non-scalar multiplications, M_{mult} is obtained. M_{dep} is the minimum required depth consumption among minimax composite polynomials that have the minimum number of non-scalar multiplications.

Second, **`DynMinDep`** puts more priority on minimizing the depth consumption rather than minimizing the number of non-scalar multiplications. The minimum depth consumption D_{dep} is obtained. D_{mult} is the minimum number of required non-scalar multiplications among minimax composite polynomials that have the minimum depth consumption.

$m_{\max}$ and $n_{\max}$ should be large enough to guarantee that the proposed algorithms find the minimax composite polynomial on R_δ that requires the minimum number of non-scalar multiplications and depth consumption among all $(\alpha-1,\delta)$-two-sided-close minimax composite polynomials on R_δ. $m_{\max}$ and $n_{\max}$ should sat-

Algorithm 6: Computation of $f(m,n,t)$ and $G(m,n,t)$ using dynamic programming

Input: $t, m_{\max}, n_{\max}$

Output: $f(m,n,t), G(m,n,t)$ for $0 \le m \le m_{\max}$ and $0 \le n \le n_{\max}$

Generate 2-dimensional table $G(m,n,t)$ for $0 \le m \le m_{\max}$ and $0 \le n \le n_{\max}$, where the components are all empty sets.

for $m \leftarrow 0$ **to** $m_{\max}$ **do**

  for $n \leftarrow 0$ **to** $n_{\max}$ **do**

    if $m \le 1$ or $n \le 1$ **then**

      $f(m,n,t) \leftarrow t$

    else

      $j \leftarrow \underset{\substack{1 \le k \\ \mathsf{mult}(2k+1) \le m \\ \mathsf{dep}(2k+1) \le n}}{\operatorname{argmax}} \mathsf{InvMinErr}(2k+1, f(m - \mathsf{mult}(2k+1), n - \mathsf{dep}(2k+1), t))$

      $f(m,n,t) \leftarrow \mathsf{InvMinErr}(2j+1, f(m - \mathsf{mult}(2j+1), n - \mathsf{dep}(2j+1), t))$

      $G(m,n,t) \leftarrow \{2j+1\} \cup G(m - \mathsf{mult}(2j+1), n - \mathsf{dep}(2j+1), t)$

    end

  end

end

isfy $f(m_{\max}, n_{\max}, 2^{1-\alpha}) \geq \delta$ and I set $m_{\max}$ and $n_{\max}$ heuristically. Note that $\mathsf{dep}(d) \leq \mathsf{mult}(d) \leq 2\mathsf{dep}(d)$ for odd d less than or equal to 31. In [24], homomorphic comparison operations were proposed for cases when $\epsilon = 2^{-\alpha}$ and $\delta = \frac{1-\epsilon}{1+\epsilon}$ and I can use the minimum number of non-scalar multiplications (or depth consumption) values as in Table 4.2 to set $m_{\max}$ and $n_{\max}$ since I also propose homomorphic comparison operations for the same case in this dissertation. Let $q(\alpha)$ be the minimum number of non-scalar multiplications (or depth consumption) for the previous algorithms. I set $m_{\max} = n_{\max} = q(\alpha)$ when minimizing the number of non-scalar multiplications and set $m_{\max} = 2q(\alpha), n_{\max} = q(\alpha)$ when minimizing the depth consumption. Then, it holds that $f(m_{\max}, n_{\max}, 2^{1-\alpha}) \geq \delta$.

M_{degs} and D_{degs} are ordered sets that store the degrees of the component minimax approximate polynomials of corresponding optimal composite polynomial when minimizing the number of non-scalar multiplications and the depth consumption, respectively. Values of $M_{\mathrm{mult}}, M_{\mathrm{dep}}$, and M_{degs} can be obtained by using Algorithm 7. Values of $D_{\mathrm{mult}}, D_{\mathrm{dep}}$, and D_{degs} can be obtained by using Algorithm 8. The procedure to find the optimal minimax composite polynomial using dynamic programming is summarized as follows:

(i) $f(m, n, t)$ and $G(m, n, t)$ are computed recursively using dynamic programming in Algorithm 6.

(ii) From the values of $f(m, n, t)$ and $G(m, n, t)$, find $M_{\mathrm{mult}}, M_{\mathrm{dep}}$, and M_{degs}, or D_{mult}, D_{dep}, and D_{degs} in Algorithms 7 and 8, respectively.

(iii) Find the component minimax approximate polynomials f_i's using modified Remez algorithm with M_{degs} or D_{degs}.

Theorem 4.3. *Let M_{mult}, M_{dep}, and M_{degs} be the output values of the* **`DynMinMult`** *algorithm in Algorithm 7 for inputs α and δ. Then, $M_{\mathrm{mult}} \leq \mathsf{MultNum}(\{f_i\}_{1 \leq i \leq k})$ for any $(\alpha - 1, \delta)$-two-sided-close composite polynomial of component polynomials with*

Algorithm 7: `DynMinMult`

Input: $\alpha, \delta, m_{\max}, n_{\max}, f(m, n, 2^{1-\alpha}), G(m, n, 2^{1-\alpha})$ for

$0 \leq m \leq m_{\max}, 0 \leq n \leq n_{\max}$

Output: M_{mult}, M_{dep}, M_{degs}

1 **for** $i \leftarrow 0$ **to** $m_{\max}$ **do**

2 **if** $f(i, n_{\max}, 2^{1-\alpha}) \geq \delta$ **then**

3 $M_{\text{mult}} \leftarrow i$

4 Go to line 7

5 **end**

6 **end**

7 **for** $j \leftarrow 0$ **to** $n_{\max}$ **do**

8 **if** $f(M_{\text{mult}}, j, 2^{1-\alpha}) \geq \delta$ **then**

9 $M_{\text{dep}} \leftarrow j$

10 Go to line 13

11 **end**

12 **end**

13 $M_{\text{degs}} \leftarrow G(M_{\text{mult}}, M_{\text{dep}}, 2^{1-\alpha})$ `// ` M_{degs}`: ordered set`

Algorithm 8: `DynMinDep`

Input: $\alpha, \delta, m_{\max}, n_{\max}, f(m, n, 2^{1-\alpha}), G(m, n, 2^{1-\alpha})$ for

$0 \leq m \leq m_{\max}, 0 \leq n \leq n_{\max}$

Output: D_{mult}, D_{dep}, D_{degs}

1 **for** $i \leftarrow 0$ **to** $n_{\max}$ **do**

2 **if** $f(m_{\max}, i, 2^{1-\alpha}) \geq \delta$ **then**

3 $D_{\text{dep}} \leftarrow i$

4 Go to line 7

5 **end**

6 **end**

7 **for** $j \leftarrow 0$ **to** $m_{\max}$ **do**

8 **if** $f(j, D_{\text{dep}}, 2^{1-\alpha}) \geq \delta$ **then**

9 $D_{\text{mult}} \leftarrow j$

10 Go to line 13

11 **end**

12 **end**

13 $D_{\text{degs}} \leftarrow G(D_{\text{mult}}, D_{\text{dep}}, 2^{1-\alpha})$ // D_{degs}: ordered set

odd degree terms $f_k \circ f_{k-1} \circ \cdots \circ f_1$. In addition, if $M_{\mathrm{mult}} = \mathsf{MultNum}(\{f_i\}_{1 \le i \le k})$, then it holds that $M_{\mathrm{dep}} \le \mathsf{DepNum}(\{f_i\}_{1 \le i \le k})$.

Proof. Let $f_k \circ f_{k-1} \circ \cdots \circ f_1$ be any $(\alpha - 1, \delta)$-two-sided-close composite polynomial of component polynomials with odd degree terms. Let $\mathsf{MultNum}(\{f_i\}_{1 \le i \le k}) = m$ and $\mathsf{DepNum}(\{f_i\}_{1 \le i \le k}) = n$. From Theorem 4.1, there is a $(\alpha - 1, \delta)$-two-sided-close minimax composite polynomial $\hat{f}_k \circ \hat{f}_{k-1} \circ \cdots \circ \hat{f}_1$ on R_δ such that $\mathsf{MultNum}(\{\hat{f}_i\}_{1 \le i \le k}) \le \mathsf{MultNum}(\{f_i\}_{1 \le i \le k})$ and $\mathsf{DepNum}(\{\hat{f}_i\}_{1 \le i \le k}) \le \mathsf{DepNum}(\{f_i\}_{1 \le i \le k})$. Assume that $m < M_{\mathrm{mult}}$. Then, $\mathsf{MultNum}(\{\hat{f}_i\}) \le \mathsf{MultNum}(\{f_i\}) = m < M_{\mathrm{mult}}$. Since $m < M_{\mathrm{mult}}$ holds and M_{mult} is the minimum i which satisfies $f(i, n_{\max}, 2^{1-\alpha}) \ge \delta$, it holds that $f(m, n_{\max}, 2^{1-\alpha}) < \delta$. Thus, there is no minimax composite polynomial $\bar{f}_k \circ \bar{f}_{k-1} \circ \cdots \circ \bar{f}_1$ on R_δ such that $\bar{f}_k \circ \bar{f}_{k-1} \circ \cdots \circ \bar{f}_1([1-\delta, 1+\delta]) \subseteq [1-2^{1-\alpha}, 1+2^{1-\alpha}]$, $\mathsf{MultNum}(\{\bar{f}_i\}_{1 \le i \le k}) \le m$, and $\mathsf{DepNum}(\{\bar{f}_i\}_{1 \le i \le k}) \le n_{\max}$. This leads to a contradiction since $\hat{f}_k \circ \hat{f}_{k-1} \circ \cdots \circ \hat{f}_1$ is a $(\alpha - 1, \delta)$-two-sided-close minimax composite polynomial on R_δ such that $\mathsf{MultNum}(\{\hat{f}_i\}_{1 \le i \le k}) \le m$ and $\mathsf{DepNum}(\{\hat{f}_i\}_{1 \le i \le k}) \le n_{\max}$.

In addition, assume that $M_{\mathrm{mult}} = m$ and $n < M_{\mathrm{dep}}$. Then, $f(m, n, 2^{1-\alpha}) < \delta$. Thus, there is no minimax composite polynomial $\bar{f}_k \circ \bar{f}_{k-1} \circ \cdots \circ \bar{f}_1$ on R_δ such that $\bar{f}_k \circ \bar{f}_{k-1} \circ \cdots \circ \bar{f}_1([1-\delta, 1+\delta]) \subseteq [1-2^{1-\alpha}, 1+2^{1-\alpha}]$, $\mathsf{MultNum}(\{\bar{f}_i\}_{1 \le i \le k}) \le m$, and $\mathsf{DepNum}(\{\bar{f}_i\}_{1 \le i \le k}) \le n$. This leads to a contradiction since $\hat{f}_k \circ \hat{f}_{k-1} \circ \cdots \circ \hat{f}_1$ is a $(\alpha - 1, \delta)$-two-sided-close minimax composite polynomial on R_δ such that $\mathsf{MultNum}(\{\hat{f}_i\}_{1 \le i \le k}) \le m$ and $\mathsf{DepNum}(\{\hat{f}_i\}_{1 \le i \le k}) \le n$.

□ □

Theorem 4.4. *Let D_{mult}, D_{dep}, and D_{degs} be the output values of the* **`DynMinDep`** *algorithm in Algorithm 8 for inputs α and δ. Then, $D_{\mathrm{dep}} \le \mathsf{DepNum}(\{f_i\}_{1 \le i \le k})$ any $(\alpha - 1, \delta)$-two-sided-close composite polynomial of component polynomials with odd degree terms $f_k \circ f_{k-1} \circ \cdots \circ f_1$. In addition, if $D_{\mathrm{dep}} = \mathsf{DepNum}(\{f_i\}_{1 \le i \le k})$, then it holds that $D_{\mathrm{mult}} \le \mathsf{MultNum}(\{f_i\}_{1 \le i \le k})$.*

Proof. The proof is omitted because the proof of Theorem 4.4 is almost the same as that of Theroem 4.3. □ □

The **MinimaxComp** algorithm that outputs an approximate value of $\mathrm{comp}(a, b)$ is now proposed as in Algorithm 9, which uses the output M_{degs} of **DynMinMult** or the output D_{degs} of **DynMinDep** algorithm. $E(a, b; d)$ and $F(a, b; d)$ are defined for the description of the **MinimaxComp** algorithm as follows.

Definition 4.8. *For $a, b \in \mathbb{R}$ and $d \in \mathbb{N}$, let $F(a, b; d)$ be the minimax approximate polynomial of degree at most d on $[-b, -a] \cup [a, b]$ for $\mathrm{sgn}(x)$ and $E(a, b; d)$ be the minimax approximation error of the minimax approximate polynomial $F(a, b; d)$.*

Algorithm 9: **MinimaxComp**

Input: $a, b \in (0, 1)$, α, ϵ

Output: An approximate value of $\mathrm{comp}(a, b)$

$\{d_1, d_2, \cdots, d_k\} \leftarrow M_{\mathrm{degs}}$ from **DynMinMult** or D_{degs} from **DynMinDep** for α and $\delta = \frac{1-\epsilon}{1+\epsilon}$

$f_1 \leftarrow F(1 - \epsilon, 1; d_1)$

$\tau_1 \leftarrow E(1 - \epsilon, 1; d_1)$

for $i \leftarrow 2$ **to** k **do**

    $f_i \leftarrow F(1 - \tau_{i-1}, 1 + \tau_{i-1}; d_i)$

    $\tau_i \leftarrow E(1 - \tau_{i-1}, 1 + \tau_{i-1}; d_i)$

end

return $\frac{f_k \circ f_{k-1} \circ \cdots \circ f_1(a-b)+1}{2}$

4.3 Numerical Results

In this section, the number of non-scalar multiplications and the depth consumption of the proposed algorithms for the approximate polynomial for the sign function are compared to those of the previous algorithm [24].

4.3.1 Computation of the Required Non-Scalar Multiplications and Depth Consumption

Let s_f and s_g be the numbers of compositions of f_n and g_n, respectively. **NewCompG** algorithm in [24] approximates $\mathrm{comp}(a, b)$ using the composite polynomial $f_n^{(s_f)} \circ g_n^{(s_g)}$ with $n = 4$. According to Lemmas 1 and 3 in [24], if $s_g \geq \lceil \frac{1}{\log 0.98 c_n^2} \cdot \log(2/\epsilon) \rceil$ and $s_f \geq \lceil \frac{1}{\log(n+1)} \cdot \log(\alpha - 2) \rceil$, then the approximation error of the output of **NewCompG**$(a, b; n, s_f, s_g)$ compared to the value of $\mathrm{comp}(a, b)$ is upper bounded by $2^{-\alpha}$, where $c_n = \frac{2n+1}{4^n}\binom{2n}{n}$.

The previous approximation method for the sign function in [24] has the best performance for $n = 4$, where the degrees of the component approximate polynomials f_n and g_n are 9, and both the required numbers of non-scalar multiplications and the depth consumption for each component polynomial are 4. Then, it should hold that $s_g \geq \lceil 0.3894 \log(2/\epsilon) \rceil$ and $s_f \geq \lceil 0.4307 \log(\alpha - 2) \rceil$.

In this dissertation, the performances of the previous and proposed algorithms are analyzed when $\epsilon = 2^{-\alpha}$, which means that the input and output of the comparison operation are required to have the same precision bits. Then both the total required numbers of non-scalar multiplications and the depth consumption are at least $4(\lceil 0.3894(\alpha + 1) \rceil + \lceil 0.4307 \log(\alpha - 2) \rceil)$.

In the proposed method, the minimum number of required non-scalar multiplications and depth consumption are computed by using the **DynMinMult** and **DynMinDep** algorithms.

4.3.2 Comparisons

Table 4.2: Comparison of the minimum number of non-scalar multiplications and the corresponding depth consumption between the previous and the proposed algorithms while minimizing the number of non-scalar multiplications.

α	number of non-scalar multiplications		depth consumption	
	the previous algorithm	the proposed algorithm	the previous algorithm	the proposed algorithm
5	16	8	16	8
6	16	11	16	10
7	24	12	24	12
8	24	14	24	14
9	24	16	24	15
10	28	18	28	16
11	28	19	28	19
12	32	20	32	20
13	32	22	32	22
14	32	24	32	23
15	36	25	36	25
16	36	27	36	26
17	40	28	40	28
18	40	30	40	29
19	40	31	40	31
20	44	33	44	32

Table 4.3: Comparison of the minimum depth consumption and the corresponding number of non-scalar multiplications between the previous and the proposed algorithms while minimizing the depth consumption.

α	number of non-scalar multiplications		depth consumption	
	the previous algorithm	the proposed algorithm	the previous algorithm	the proposed algorithm
5	16	10	16	7
6	16	14	16	8
7	24	14	24	10
8	24	18	24	11
9	24	18	24	13
10	28	21	28	14
11	28	25	28	15
12	32	28	32	16
13	32	31	32	17
14	32	31	32	19
15	36	34	36	20
16	36	37	36	21
17	40	40	40	22
18	40	43	40	23
19	40	47	40	24
20	44	50	44	25

Table 4.4: The ordered sets M_{degs} and D_{degs} that store the degrees of the optimal component minimax approximate polynomials in **`DynMinMult`** and **`DynMinDep`** algorithms, respectively.

α	M_{degs}	D_{degs}
5	$\{9, 9\}$	$\{7, 13\}$
6	$\{5, 7, 9\}$	$\{15, 15\}$
7	$\{9, 9, 9\}$	$\{7, 7, 13\}$
8	$\{3, 9, 9, 9\}$	$\{7, 15, 15\}$
9	$\{7, 9, 9, 9\}$	$\{7, 7, 7, 13\}$
10	$\{3, 7, 7, 9, 9\}$	$\{7, 7, 13, 15\}$
11	$\{5, 9, 9, 9, 9\}$	$\{7, 7, 15, 31\}$
12	$\{9, 9, 9, 9, 9\}$	$\{7, 15, 15, 31\}$
13	$\{3, 9, 9, 9, 9, 9\}$	$\{15, 15, 15, 31\}$
14	$\{7, 9, 9, 9, 9, 9\}$	$\{7, 7, 13, 15, 31\}$
15	$\{3, 5, 9, 9, 9, 9, 9\}$	$\{13, 7, 15, 15, 31\}$
16	$\{5, 7, 9, 9, 9, 9, 9\}$	$\{13, 15, 15, 15, 31\}$
17	$\{9, 9, 9, 9, 9, 9, 9\}$	$\{13, 15, 15, 31, 31\}$
18	$\{7, 3, 9, 9, 9, 9, 9, 9\}$	$\{13, 15, 31, 31, 31\}$
19	$\{5, 9, 9, 9, 9, 9, 9, 9\}$	$\{15, 31, 31, 31, 31\}$
20	$\{9, 9, 9, 9, 9, 9, 9, 11\}$	$\{31, 31, 31, 31, 31\}$

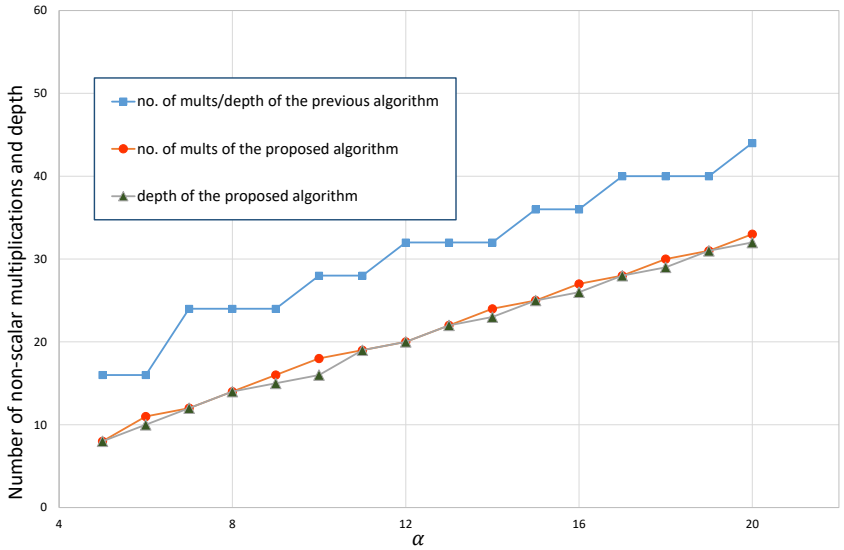

Figure 4.1: Comparison of the minimum number of non-scalar multiplications and the corresponding depth consumption between the previous and the proposed algorithms while minimizing the number of non-scalar multiplications.

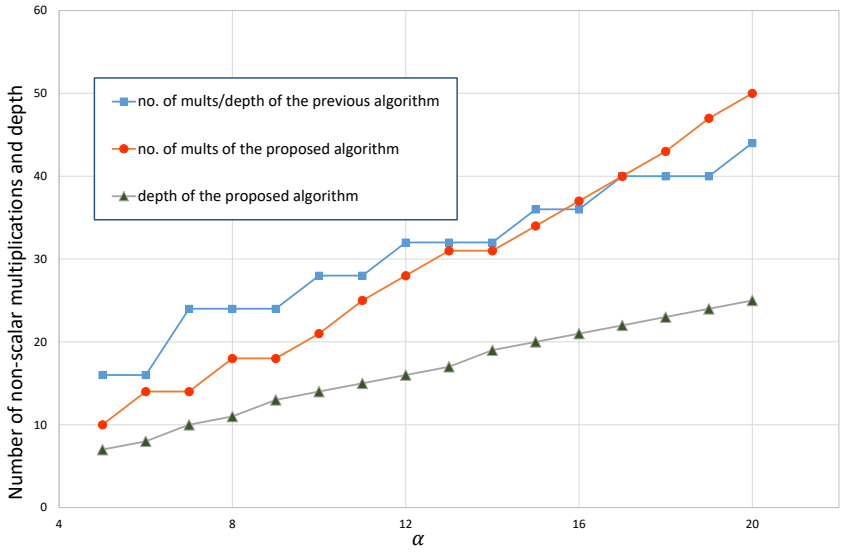

Figure 4.2: Comparison of the minimum depth consumption and the corresponding number of non-scalar multiplications between the previous and the proposed algorithms while minimizing the depth consumption.

Table 4.2 shows the comparison of the minimum number of non-scalar multiplications and the corresponding depth consumption between the previous algorithm and the proposed algorithm **DynMinMult** while minimizing the number of non-scalar multiplications. It can be seen from Table 4.2 that the minimum number of the required non-scalar multiplications and the corresponding depth consumption for the proposed algorithms are reduced by about 33% and 35% on average, respectively, compared to those of the previous algorithm. The proposed algorithm **DynMinMult** intends to minimize the number of non-scalar multiplications, however, the corresponding depth consumption is also decreased. Figure 4.1 describes Table 4.2 as a graph.

Table 4.3 shows the comparison of the minimum depth consumption and the corresponding number of non-scalar multiplications between the previous algorithm and the proposed algorithm **DynMinDep** while minimizing the depth consumption. It can be seen from Table 4.3 that the non-scalar multiplications and the corresponding depth consumption for the proposed algorithms are reduced by about 10% and 47% on average, respectively, compared to those of the previous algorithm. If $\alpha \geq 16$, then the number of non-scalar multiplications for the proposed algorithm is slightly larger than that for the previous algorithm. However, when bootstrapping is used, the proposed algorithm requires lower time complexity than the previous algorithm since bootstrapping due to large depth consumption requires higher time complexity than non-scalar multiplication operations. Figure 4.2 describes Table 4.3 as a graph.

Table 4.4 shows the ordered sets M_{degs} and D_{degs} that store the degrees of the optimal component minimax approximate polynomials when minimizing the number of non-scalar multiplications and depth consumption, respectively.

Chapter 5

Conclusions

In this dissertation, modification of FrodoKEM using Gray and error-correcting codes and optimal composition of approximate component polynomials with odd degree terms for homomorphic comparison operation were proposed.

First, FrodoPKE is viewed as a digital communication system. BCH code parameters were designed for FrodoKEM and Gray coding is also used in FrodoKEM to decrease the DFR. It was proved that the modified FrodoKEM achieves IND-CCA security and the performance of the modified FrodoKEM was analyzed. First, it was shown that the security level of FrodoKEM can be increased for the modified FrodoKEM. Second, it was found that the message size of Frodo-640 is increased. Finally, it was shown that the bandwidth of Frodo-640 is reduced. I confirmed through numerical analysis that the DFR is very high when Gray coding is not used.

Second, I proposed a new approximation method for the homomorphic comparison operation using minimax composite polynomials obtained by the modified Remez algorithm. My main idea is to find the minimax composite polynomial on R_δ that requires the minimum number of non-scalar multiplications and depth consumption among all $(\alpha - 1, \delta)$-two-sided-close minimax composite polynomials on R_δ. It was proved that the obtained minimax composite polynomial on R_δ requires less number of non-scalar multiplications and depth consumption than any $(\alpha - 1, \delta)$-two-

sided-close composite polynomial of component polynomials with odd degree terms. Since the brute-force search requires considerable time for α, I proposed polynomial-time algorithms that obtain the best minimax composite polynomials by using dynamic programming. It can be seen from numerical analysis that when the number of non-scalar multiplications is minimized, the minimum number of required non-scalar multiplications and the corresponding depth consumption for the proposed algorithm **`DynMinMult`** are reduced by about 33% and 35% on average, respectively, compared to those for the previous algorithm. In addition, when the depth consumption is minimized, the minimum number of required non-scalar multiplications and the corresponding depth consumption for the proposed algorithm **`DynMinDep`** are reduced by about 10% and 47% on average, respectively, compared to those for the previous algorithm.

Bibliography

[1] M. Naehrig, E. Alkim, J. Bos, L. Ducas, K. Easterbrook, B. LaMacchia, P. Longa, I. Mironov, V. Nikolaenko, C. Peikert, A. Raghunathan, and D. Stebila, "FrodoKEM," Technical report, Available: https://csrc.nist.gov/projects/post-quantum-cryptography/round-1-submissions.

[2] T. Poppelmann, E. Alkim, R. Avanzi, J. Bos, L. Ducas, A. Piedra, P. Schwabe, and D. Stebila, "NewHope," Technical report, Available: https://csrc.nist.gov/projects/post-quantum-cryptography/round-1-submissions.

[3] P. Schwabe, R. Avanzi, J. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. M. Schanck, G. Seiler, and D. Stehle, "CRYSTALS-KYBER," Technical report, Available: https://csrc.nist.gov/projects /post-quantum-cryptography/round-1-submissions.

[4] D. Micciancio, "Lattice-based cryptography," *Post-Quantum Cryptography*, LNCS, Berlin, Germany: Springer, 2011, pp. 147–191.

[5] C. Peikert, "A decade of lattice cryptography," *Foundations and Trends in Theoretical Computer Science*, vol. 10, no. 4, pp. 283–424, 2016.

[6] O. Regev, "On lattices, learning with errors, random linear codes, and cryptography," *Journal of the ACM*, vol. 56, no. 6, pp. 1–37, 2009.

[7] P. W. Shor, “Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer,” *SIAM Review*, vol. 41, no. 2, pp. 303–332, 1999.

[8] S. Bhattacharya, O. Garcia-Morchon, T. Laarhoven, R. Rietman, M. O. Saarinen, L. Tolhuizen, and Z. Zhang, “Round5: compact and fast post-quantum public-key encryption,” Cryptol. ePrint Arch., Tech. Rep. 2019/090, 2019. [Online]. Available: https://eprint.iacr.org/2019/090.

[9] M. O. Saarinen, “HILA5: On reliability, reconciliation, and error correction for Ring-LWE encryption,” in *Proc. International Conference on Selected Areas in Cryptography*, LNCS, vol. 10719. Berlin, Germany: Springer, 2017, pp. 192–212.

[10] O. Garcia-Morchon, Z. Zhang, S. Bhattacharya, R. Rietman, L. Tolhuizen, J. Torre-Arce, and H. Baan, “Round2,” Technical report, Available: https://csrc.nist.gov/projects/post-quantum-cryptography/round-1-submissions.

[11] X. Lu, Y. Liu, D. Jia, H. Xue, J. He, and Z. Zhang, “LAC,” Technical report, Available: https://csrc.nist.gov/projects/post-quantum-cryptography/round-1-submissions.

[12] X. Lu, Y. Liu, Z. Zhang, D. Jia, H. Xue, J. He, and B. Li, “LAC: Practical ring-LWE based public-key encryption with byte-level modulus*,” Cryptol. ePrint Arch., Tech. Rep. 2018/1009, 2018. [Online]. Available: https://eprint.iacr.org/2018/1009.

[13] M. Hamburg, “Three Bears,” Technical report, Available: https://csrc.nist.gov/projects/post-quantum-cryptography/round-1-submissions.

[14] Y. Zhao, Z. Jin, B. Gong, and G. Sui, “KCL,” Technical report, Available: https://csrc.nist.gov/projects/post-quantum-cryptography/round-1-submissions.

[15] E. Alkim, L. Ducas, T. Poppelmann, and P. Schwabe, "Post-quantum key exchange-a NewHope," in *Proc. 25th USENIX Security Security '16*, Santa Clara, CA, USA, 2016, pp. 327–343.

[16] T. Fritzmann, T. Poppelmann, and J. Sepulveda, "Analysis of error-correcting codes for lattice-based key exchange," in *Proc. International Conference on Selected Areas in Cryptography*, LNCS, Berlin, Germany: Springer, 2018, pp. 369–390.

[17] P. Martins, L. Sousa, and A. Mariano, "A survey on fully homomorphic encryption: an engineering perspective," *ACM Computing Surveys (CSUR)*, vol. 50, no. 6, pp. 1–33, 2017.

[18] C. Gentry, "Fully homomorphic encryption using ideal lattices," in *Proc. of the Forty-First Annual ACM Symposium on Theory of Computing*, 2019, pp. 169–178.

[19] Z. Brakerski, C. Gentry, and V. Vaikuntanathan, "(Leveled) Fully homomorphic encryption without bootstrapping," *ACM Transactions on Computation Theory (TOCT)*, vol. 6, no. 3, pp. 1–36, 2014.

[20] J. Fan and F. Vercautern, "Somewhat practical fully homomorphic encryption," Cryptol. ePrint Arch., Tech. Rep. 2012/144, 2012. [Online]. Available: https://eprint.iacr.org/2012/144.

[21] I. Chillotti, N. Gama, M. Georgieva, and M. Izabachene, "TFHE: fast fully homomorphic encryption over the torus," *Journal of Cryptology*, vol. 33, no. 1, pp. 34–91, 2020.

[22] J. Cheon, A. Kim, M. Kim, and Y. Song, "Homomorphic encryption for arithmetic of approximate numbers," *Proc. International Conference on the Theory and Application of Cryptology and Information Security*, LNCS, Berlin, Germany: Springer, 2017, pp. 409–437.

[23] J. Cheon, D. Kim, D. Kim, H. Lee, and K. Lee, "Numerical method for comparison on homomorphically encrypted numbers," in *Proc. International Conference on the Theory and Application of Cryptology and Information Security*, LNCS, Berlin, Germany: Springer, 2019, pp. 415-445.

[24] J. Cheon, D. Kim, and D. Kim, "Efficient homomorphic comparison methods with optimal complexity," Cryptol. ePrint Arch., Tech. Rep. 2019/1234, 2019. [Online]. Available: https://eprint.iacr.org/2019/1234.

[25] D. Hofheinz, K. Hovelmanns, and E. Kiltz, "A modular analysis of the Fujisaki-Okamoto transformation," *Proc. TCC 2017: 15th Theory of Cryptography Conference*, LNCS, Berlin, Germany: Springer, 2017, pp. 341–371.

[26] J. P. D'Anvers, F. Vercauteren, and I. Verbauwhede, "On the impact of decryption failures on the security of LWE/LWR based schemes," Cryptol. ePrint Arch., Tech. Rep. 2018/1089, 2018. [Online]. Available: https://eprint.iacr.org/2018/1089.

[27] R. C. Bose and D. K. Ray-Chaudhuri, "On a class of error-correcting binary group codes," *Information and Control*, vol. 3, no. 1, pp. 68–79, 1960.

[28] D. Gorenstein, W. W. Peterson, and N. Zierler, "Two-error correcting Bose-Chaudhuri codes are quasi-perfect," *Information and Control*, vol. 3, no. 3, pp. 291–294, 1960.

[29] M. Walters and S. S. Roy, "Constant-time BCH error-correcting code," Cryptol. ePrint Arch., Tech. Rep. 2019/155, 2019. [Online]. Available: https://eprint.iacr.org/2019/155.

[30] E. W. Cheney, *Introduction to Approximation Theory*. Cambridge, U.K.:McGraw-Hill, 1966.

[31] J. Lee, E. Lee, Y. Lee, Y. Kim, and J. No, "Optimal minimax polynomial approximation of modular reduction for bootstrapping of approximate homomorphic encryption," Cryptol. ePrint Arch., Tech. Rep. 2020/552, 2020. [Online]. Available: https://eprint.iacr.org/2020/552.

[32] Y. Lee, J. Lee, Y. Kim, and J. No, "Near-optimal polynomial for modulus reduction using l2-norm for approximate homomorphic encryption," Cryptol. ePrint Arch., Tech. Rep. 2020/488, 2020. [Online]. Available: https://eprint.iacr.org/2020/488.

[33] E. Y. Remez, "Sur la determination des polynomes d'approximation de degre donnee," *Comm. Soc. Math. Kharkov*, vol. 10, no. 196, pp. 41–63, 1934.

[34] V. Lyubashevsky, C. Peikert, and O. Regev, "On ideal lattices and learning with errors over rings," in *Proc. Advances in Cryptology–EUROCRYPT*, LNCS, vol. 6110. Berlin, Germany: Springer, 2010, pp. 1–23.

[35] J. Bos, C. Costello, L. Ducas, L. Mironov, M. Naehrig, V. Nikolaenko, A. Raghunathan, and D. Stebila, "Frodo: Take off the ring! practical, quantum-secure key exchange from LWE," in *Proc. the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 2016, pp. 1006–1018.

[36] R. Lindner and C. Peikert, "Better key sizes (and attacks) for LWE-based encryption," in *Proc. Cryptographers' Track at the RSA Conference*, LNCS, vol. 6558, 2011, pp. 319–339.

[37] J. G. Proakis and M. Salehi, *Communication Systems Engineering*, 2nd ed. NJ, USA: Prentice Hall, 1994.

[38] Y. Cassuto, M. Schwartz, V. Bohossian, and J. Bruck, "Codes for multi-level flash memories: correcting asymmetric limited-magnitude errors," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Nice, France, Jun. 2007, pp. 1176–1180.

[39] H. Helgert and R. Stinaff, "Shortened BCH codes," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 818–820, 1973.

[40] R. E. Blahut, *Algebraic Codes for Data Transmission*. Cambridge, U.K.: Cambridge university press, 2003.

[41] J. P. D'Anvers, F. Vercauteren, and I. Verbauwhede, "The impact of error dependencies on Ring/Mod-LWE/LWR based schemes," in *Proc. 2019 International Conference on Post-Quantum Cryptography*, LNCS, Berlin, Germany: Springer, 2019, pp. 225–246.

[42] R. Gilad-Bachrach, N. Dowlin, K. Laine, K. Lauter, M. Naehrig, and J. Wernsing, "Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy," in *Proc. International Conference on Machine Learning*, 2016, pp. 201–210.

[43] C. Juvekar, V. Vaikuntanathan, and A. Chandrakasan, "GAZELLE: A low latency framework for secure neural network inference," in *Proc. 27th USENIX Secur. Symp. (USENIX Security)*, 2018, pp. 1651–1669.

[44] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. on Pattern Analysis & Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[45] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001.

[46] C. Boura, N. Gama, and M. Georgieva, "Chimera: a unified framework for B/FV, TFHE and HEAAN fully homomorphic encryption and predictions for deep learning," Cryptol. ePrint Arch., Tech. Rep. 2018/758, 2018. [Online]. Available: https://eprint.iacr.org/2018/758.

[47] D. Chialva and A. Dooms, “Conditionals in homomorphic encryption and machine learning applications,” Cryptol. ePrint Arch., Tech. Rep. 2018/1032, 2018. [Online]. Available: https://eprint.iacr.org/2018/1032.

[48] R. E. Goldschmidt, *Applications of Division by Convergence*. PhD thesis, Massachusetts Institute of Technology, 1964.

[49] M. S. Paterson and L. J. Stockmeyer, “On the number of nonscalar multiplications necessary to evaluate polynomials,” *SIAM Journal on Computing*, vol. 2, pp. 60–66, 1973.

초 록

이 학위 논문에서는, 다음 두 가지 내용이 연구되었다.

(i) FrodoKEM을 그레이 부호 및 오류정정부호를 사용하여 개선

(ii) 동형 비교 연산을 위해 합성 다항식을 사용한 부호 함수의 최적 미니맥스 다항식 근사

먼저, 그레이 부호 및 오류정정부호를 사용하여 FrodoKEM을 변형시키는 방법이 연구되었다. 격자기반암호는 가장 유망한 포스트 양자 암호 스킴이다. 많은 격자기반암호 시스템 중에서 FrodoKEM은 learning with errors (LWE) 문제에 기반을 둔 잘 알려진 키-캡슐화 메커니즘 (KEM) 이며 구조를 갖지 않은 격자 문제에 기반을 둔 어려움을 가진다는 장점이 있다. NIST 포스트 양자 암호 표준화 라운드 2에 발표된 LAC, Three Bears, Round5와 같이 성능 개선을 위해 오류정정부호를 사용하는 많은 암호 시스템들이 있다. 그러나 FrodoKEM과 같이 링 구조를 사용하지 않는 격자기반 암호 시스템에서는 전송되는 심볼 개수가 작기 때문에 오류정정부호를 사용하기 어렵다. 나는 암호화된 심볼로부터 변환된 비트들을 부호화하여 오류정정부호와 그레이 부호를 FrodoKEM에 적용하는 방법을 제안하였다. 제안한 알고리즘은 FrodoKEM의 보안성 레벨 혹은 데이터전송량을 향상하고 기존 128비트보다 50% 많은 192비트가 변형된 Frodo-640에서 전송될 수 있음을 보여주었다.

두 번째로, 합성 다항식을 사용한 부호 함수의 최적 미니맥스 다항식 근사가 연구되었다. 두 숫자의 비교 함수는 딥러닝 및 데이터 처리 시스템을 포함한 많은 응용에서 가장 많이 사용되는 연산 중 하나이다. 암호문 상에서의 덧셈과 곱셈만

지원하는 동형 암호에서 비교 함수를 효율적으로 계산하는 몇몇 연구가 진행되었다. 동형 암호에서 합성 다항식을 사용하여 부호 함수를 근사하는 비교 방법은 동형 비교 연산이라고 불리는데 최근 새로운 동형 비교 연산 방법이 제안되었고 그 방법이 최적 점근적 복잡도를 가진다는 것이 증명되었다. 본 논문에서 나는 미니맥스 근사다항식의 합성함수를 사용하여 동형암호에서 부호 함수를 근사하는 새로운 최적 알고리즘을 제안한다. 미니맥스 근사 다항식은 modified Remez 알고리즘에 의해 얻을 수 있다. 제안하는 알고리즘은 임의의 부호 함수를 근사하는 홀수 차수 항들을 가진 다항식의 합성 다항식을 사용하는 방법보다 더 적은 넌스칼라 곱 및 뎁스 소모를 사용한다는 것이 증명되었다. 또한, 제안한 동형 비교 연산에 대한 다이나믹 프로그래밍을 사용한 최적 다항시간 알고리즘이 제안되었다. 수치 분석 결과, 넌스칼라 곱 개수를 최소로 할 때, 제안하는 알고리즘은 필요한 넌스칼라 곱 개수와 뎁스 소모를 기존 방법의 필요한 넌스칼라 곱 개수 및 뎁스 소모보다 각각 33%, 35%정도 감소시킨다. 또한, 뎁스 소모를 최소로 할 때, 제안하는 알고리즘은 필요한 넌스칼라 곱 개수와 뎁스 소모를 기존 방법의 필요한 넌스칼라 곱 개수 및 뎁스 소모보다 각각 10%, 47%정도 감소시킨다.

주요어: 격자기반암호, 그레이 부호, 동형 비교 연산, 미니맥스 근사, 부호 함수, 오류정정부호, 완전동형암호, 포스트 양자 암호, FrodoKEM, Remez 알고리즘

학번: 2014-22573

감사의 글

여호와는 나의 목자시니 내가 부족함이 없으리로다. 그가 나를 푸른 초장에 누이시며 쉴 만한 물가으로 인도하시는도다. (시 23:1-2)

지금까지 늘 저를 사랑하시고 선하게 인도해 주신 하나님께 감사드립니다. 전문연구요원 준비, 프로젝트 업무, 연구 등에 있어서 많은 난관이 있었지만, 하나님의 인도하심이 있었기에 그 모든 것을 극복하고 여기까지 올 수 있었습니다.

연구자로서, 교육자로서 뛰어난 본을 보여 주시면서 헌신적으로 지도해 주신 노종선 교수님께 진심으로 감사드립니다. 많은 연구 아이디어와 조언을 주시고 세세하고 꼼꼼하게 논문을 검토해 주신 것이 제게 큰 힘이 되었습니다.

많은 연구 조언을 해주시고 프로젝트 업무를 지도해 주신 조선대 김영식 교수님, 연구 수행과 논문 작성에 많은 도움을 주신 한양대 신동준 교수님, 시뮬레이션을 도와주신 희열형, 방장 일과 과제 업무를 훌륭하게 수행하여 연구실을 책임지면서 늘 힘이 되어준 용우, 핵심적인 연구 아이디어를 제공하고 연구 결과를 수학적으로 증명하는 데 큰 도움을 준 준우, 그 밖에도 많은 조언과 도움을 주신 부호 및 암호 연구실의 모든 분께 진심으로 감사의 말씀을 드리고 싶습니다.

믿음의 삶을 본으로 보여주시고 삶의 이정표가 되어주신 하선생님 내외분께 진심으로 감사를 드립니다. 무슨 일에든지 믿음을 배울 수 있도록 격려와 조언을 아끼지 않으신 김상완형제님, 권혁준형제님, 양철원형제님, 상혁형과 늘 기도로 함께해 주신 용하형, 성익형, 경환형, 야마다형, 승운형, 영기, 명현, 원, 진욱, 회영 모두에게 감사의 마음을 전합니다.

끝으로, 저를 길러 주시고 늘 사랑해 주시고 은혜를 베풀어 주신 부모님과 기도로 후원해 주시는 할아버지, 할머니, 누나 모두에게 감사의 말씀을 드리고 싶습니다. 하나님의 은혜와 수많은 분의 든든한 후원 속에 무사히 졸업할 수 있었습니다. 저를 도와주신 모든 분께 감사의 마음을 담아 이 논문을 바칩니다.