University of Limoges

Securing Online Analytics of Shared Data Warehouses

Student:
DOAN Thi Van Thao

Supervisors:
Mohamed-Lamine MESSAI
Gérald GAVIN
Jérome DARMONT

A thesis submitted in fulfillment of the requirements for the degree of Master CRYPTIS

in the

ERIC laboratory
Lumière University Lyon 2

August 31, 2022

Acknowledgements

The research depicted in this work is funded by the French National Research Agency (ANR), project ANR-19-CE23-0005 BI4people (Business Intelligence for the people), ERIC laboratory (**E**ntrepôts, **R**eprésentation et **I**ngénierie des **C**onnaissances).

In the study for and preparation of the thesis, I am extremely grateful to many people, who have helped me to make this work possible.

First and foremost, I would like to render my warmest thanks to my main supervisor, Professor Mohamed-Lamine Messai for his friendly guidance and expert advice throughout the whole process. I am also very grateful to my two co-supervisors in the project, Professor Gérald Gavin and Professor Jérome Darmont for giving me the valuable opportunity to do a research in this promising cryptographic topic, and with some of the most talented people.

I would also wish to express my gratitude to Professor Duong-Hieu Phan and the Vingroup Graduate Scholarship. With their generous support, I had a great experience of studying the master's degree in Cryptography and Information Security (CRYPTIS), one of the key areas at an excellent university in France.

I am also deeply thankful to all of my lecturers at the Faculty of Science and Technology, University of Limoges. Thank you for providing me with a wide range of in-depth cryptography and security courses. These hands-on lessons would assist me with necessary equipment to achieve my goal of becoming an expert in the information security sector, as well as leave me well prepared for any postgraduate endeavours in the future.

I would sincerely like to thank all my cherished Vietnamese friends in Limoges who were with me and supported me through thick and thin. I still remember the moments when we did projects together and how cheerful they were after getting the job done. I am also grateful to my colleagues and professors at ERIC laboratory, who provided me with a professional working environment and practical knowledge before doing a PhD.

My special thanks are due to my beloved family for their constant encouragement, along with their continuous understanding and inspiration, that made this project possible.

Abstract

DOAN Thi Van Thao

Securing Online Analytics of Shared Data Warehouses

With the increased need for data confidentiality in different applications of our daily life, homomorphic encryption (HE) arises as a promising cryptographic topic. HE allows to perform computations directly on encrypted data (ciphertexts) without decryption in advance. Since the results of calculations remain encrypted and can only be decrypted by the data owner, confidentiality is warranted and any third party can operate on ciphertexts without access to decrypted data (plaintetxts). Applying a homomorphic cryptosystem in a real-world application depends on its resource efficiency. Several works compared different HE schemes and gave the stakes of this research field. However, the existing articles either do not deal with recently proposed HE schemes (such as CKKS) or focus only on one type of HE.

Within the context of the BI4people project (Business intelligence for the people), our objective is to conduct an extensive comparison and evaluation of homomorphic cryptosystems' performance based on their experimental results. This evaluation aims to explore, analyze, and compare existing solutions to data confidentiality problem using Homomorphic Encryption, followed by proposing the most appropriate and sufficient solution for BI4people. Moreover, the obtained results of performance evaluation can also be used to a variety of other situations. The study covers all three families of HE, including several notable schemes such as BFV, BGV, CKKS, RSA, El-Gamal, and Paillier. In addition, we present the basics of the HE schemes, in a great part of their principles and mathematical models, as well as their implementation specification in most-used HE libraries, namely Microsoft SEAL, PALISADE, and HElib.

Contents

List of Figures

List of Tables

List of Abbreviations

HE	**H**omomorphic **E**ncryption
PHE	**P**artially **H**omomorphic **E**ncryption
SWHE	**S**ome**W**hat **H**omomorphic **E**ncryption
FHE	**F**ully **H**omomorphic **E**ncryption
BFV	**B**rakerski/**F**an-**V**ercauteren
BGV	**B**rakerski-**G**entry-**V**aikuntanathan
CKKS	**C**heon-**K**im-**K**im-**S**ong
HEAAN	**H**omomorphic **E**ncryption for **A**rithmetic of **A**pproximate **N**umbers
IFP	**I**nteger **F**actorization **P**roblem
DLP	**D**iscrete **L**ogarithmic **P**roblem
IND-CPA	**IND**istinguishability under **C**hosen-**P**laintext **A**ttack
IND-CCA	**IND**istinguishability under **C**hosen-**C**iphertext **A**ttack

Chapter 1

Introduction

For a long time, security has always been a controversial topic due to its importance in technology particularly and in society generally. When implementing a technological tool or service, the first and foremost concern of researchers is about the applicable security that it can provide.

By the dramatically inevitable data growth of nearly all organizations, the demand for data storage and computation has been increasing significantly in the last decades. A traditional infrastructure for data management, such as in-house or local services, can provide only a limited storage and access controls. In the Internet-based world, this method is no longer applicable with a huge amount of sensitive data being generated every second from business transactions. One potential solution for this problem is to seek a third-party expert outside of the company to place its trust. However, to apply this model, we need to face one of its biggest challenges: *data confidentiality*.

In this situation, cryptography has come to the forefront to provide both data confidentiality and computation for this outsourcing problem. As the foundation of modern security systems, cryptography helps to ease the concern of data leakage to an untrusted third party or server side. An user must now encrypt the data before sending it to the server. Later, after retrieving encrypted data from the server, only the user can decrypt it using his secret key and get its value. Although this technique would preserve the data privacy, the encrypted data is not meaningful for the server, so it is not able to maintain its computation's efficiency. That was why for the moment, a new cryptographic topic, called Homomorphic Encryption, got a major attention when it allows to perform certain computable functions on the encrypted data while preserving the features of the function and format of the ciphertexts. In [Aca+18], A. Acar and his colleagues beautifully illustrated this process on Figure 1.1, where C is a Client and S is a Server.

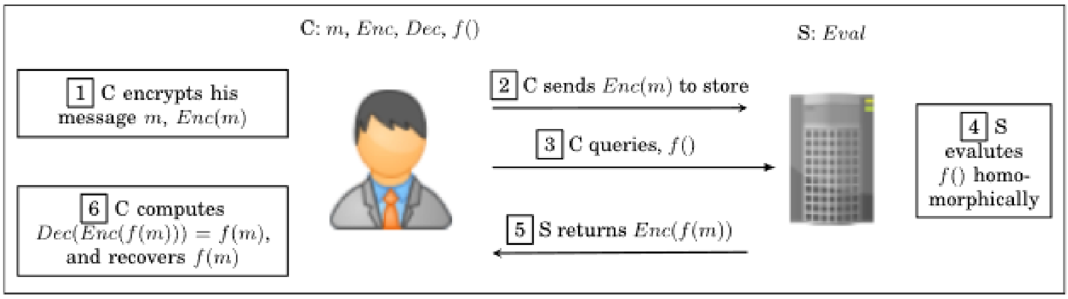

FIGURE 1.1: A simple client-server HE scenario.

Also, following this survey, HE can be categorized under three types of schemes with respect to the number of allowed operations on the encrypted data: (1) *Partially Homomorphic Encryption* (PHE) allows only one type of operation with an unlimited number of times. (2) *Somewhat Homomorphic Encryption* (SWHE) allows some types of operations with a limited number of times. (3) *Fully Homomorphic Encryption* (FHE) allows an unlimited number of operations for an unlimited number of times. Figure 1.2 presents the most known HE-based systems and their timeline, while their application scenarios are demonstrated in Table 1.1.

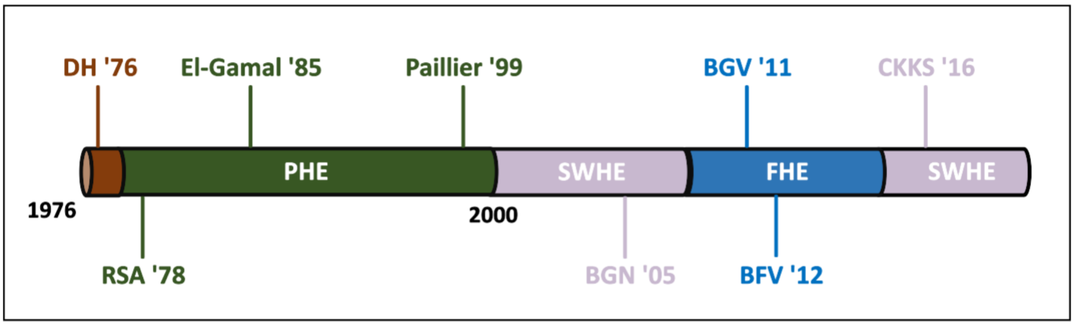

FIGURE 1.2: Timeline of several important HE schemes.

So far, there are many HE schemes that have been introduced. Within the scope of the project, we focus on the ones which are the most widely used in cryptography applications and are the basis for other schemes. Being introduced in 1978, RSA is one of the first public-key encryption methods for securing communication on the Internet, inspired by Diffie-Hellman's research ([DH19], 1976). Little while later, El-Gamal ([ElG85], 1985) and Paillier cryptosystems ([Pai99], 1999) were proposed respectively, marking an important milestone for PHE. The calculation on ciphertexts remained limited until 2009, when C. Gentry presented a FHE scheme ([Gen09], 2009). Three years later, based on the Gentry's work, two main FHE schemes to perform exact computations over finite fields and integers were born, Brakerski-Gentry-Vaikuntanathan (BGV) [BGV14] and Brakerski/ Fan-Vercauteren (BFV) [FV12]. The latest newcomer to join SWHE recently is CKKS ([Che+17], 2016), which allows to perform computations over approximated numbers. CKKS is an indispensable element of HE family, where it complements previous schemes by natively dealing with real and complex numbers.

Scheme	Application
RSA, 1978 [RSA78]	Banking and credit card transaction (Parmar et al., [Par+14])
ElGamal, 1985 [ElG85]	In Hybrid Systems (Parmar et al., [Par+14])
Paillier, 1999 [Pai99]	E-Voting (Parmar et al., [Par+14])
BGN, 2005 [BGN05]	A Novel IoT Data Protection Scheme Based on BGN Cryptosystem (S. Halder et al., [HC21])
BGV, 2011 [BGV14]	For the Security of Integer Polynomials (Parmar et al., [Par+14])
BFV, 2012 [FV12]	A fast oblivious linear evaluation (OLE) protocol (Leo de Castro [De 20])
CKKS, 2016 [Che+17]	Homomorphic Machine Learning Big Data Pipeline for the Financial Services Sector (Masters et al., [Mas+19])

TABLE 1.1: HE schemes and their applications.

Before moving on to the following section, Figure 1.3 shows how HE is applied to the BI4people context. In the project, there are two main participants, a server and a user. The local system belongs to the user's part. Depending on whether data confidentiality is required or not, the context is divided into two different possibilities. In case the data is not very sensible or important, the user has the option to send data to the server without encryption. On the other hand, he/she can select to encrypt data before sending them. In this situation, he/she will be prompted to choose between encrypting full data (all of his data) or partial data (a part of them). In both cases, he/she goes forward by selecting needed operations on his/her data, such as addition, multiplication, etc. After that, the local system takes charge by automatically generating key pair, encrypting required data, and sending them, together with the public key to the server. The server then does necessarily required computations and returns the results. Using secrete key owned by the user only, the local system will decrypt received results and give the user access to the values in plaintext.

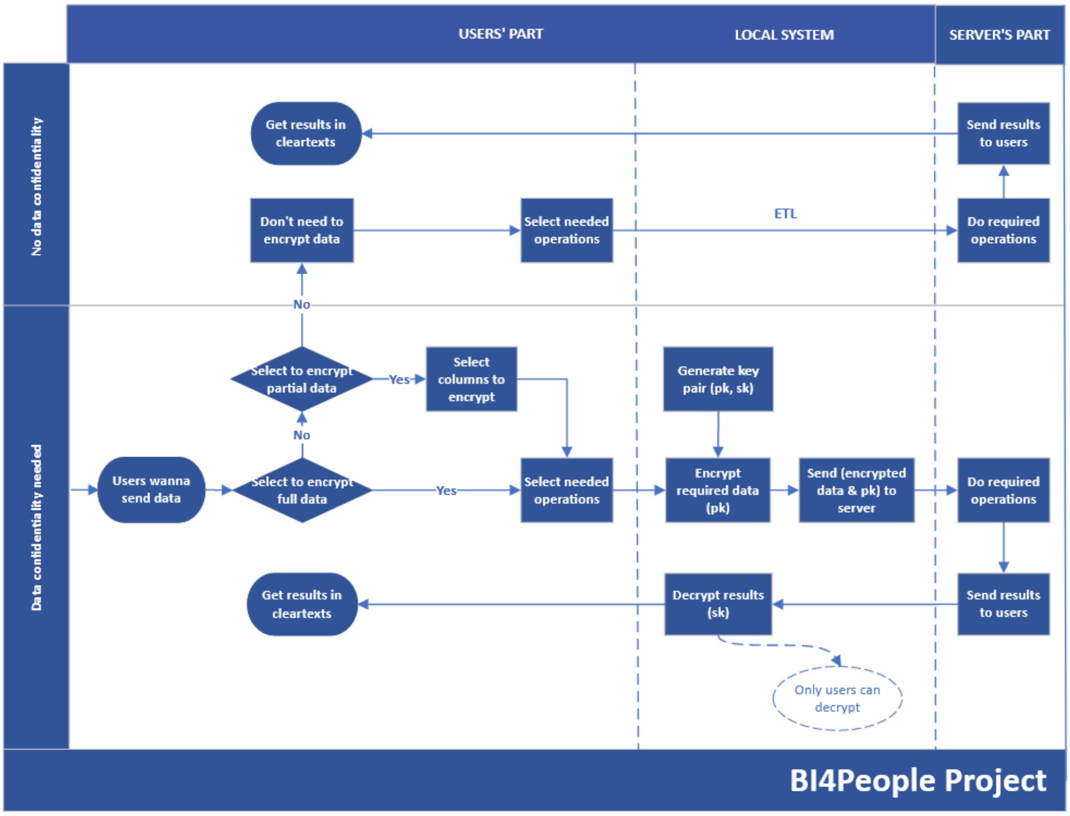

FIGURE 1.3: HE application in BI4people project.

Organization of the report

The rest of the thesis is structured as follows: chapter 2 reviews some of the related works in the similar field, regarding the performance evaluation of different HE schemes. The most important properties of HE schemes and their libraries are discussed in chapter 3. Then, chapter 4 elaborates the implementation method and results of evaluation analysis. In chapter 5, a discussion on the security of HE under notable security notions and Shor's quantum algorithm is given. Finally, chapter 6 presents a conclusion and indicates directions of future work.

Chapter 2

Previous Work in the Field

As mentioned above, one of our related works is a survey conducted by A. Acar et al. in 2018 [Aca+18], which covers important PHE, SWHE, and all the main FHE schemes. Similarly, in 2017 the survey of P. Martins et al. [MSM17] presents fundamental concepts of FHE schemes and their performance, mainly from an engineering perspective, refraining from introducing complex mathematical definitions. However, the articles do not mention CKKS encryption [Che+17], an usefully practical HE proposed recently in 2016, which allows to compute real and complex input numbers. Lately, a study of Kim et al. [KPZ21], published in 2021, implements their improved variants of BFV and BGV in PALISADE and evaluate their experimental performance for several benchmark computations. From a same point of view, Lepoint and Naehrig in [LN14] offers theoretical and practical comparisons of different HE schemes, as well as explains how to choose parameters to ensure algorithms' correctness and security. Even so, the papers delve deeply into the mathematics, which is more suitable for expert readers and mathematicians. In contrast, the survey conducted by Alaya et al. in [ALM20] makes a easy-to-understand comparison of advantages and limitations of different HE algorithms. Unfortunately, it only presents the theoretical information of the schemes, while implementation aspect has not been brought up. Most recently, Sidorov and his colleagues [SWN22] publish an article relating to performance analysis of HEs in several libraries, but the paper does not specify which homomorphic schemes were used in each libraries, either the input parameters. In opposition to [SWN22], Migliore et al. [MBF16] proposes a study of the current best solutions for setting up parameters of HE schemes, but only approaches to SWHEs.

Considering the related works and their scopes summarized in Table 2.1, it is obvious that among existing HE surveys, they either do not study newborn schemes (such as CKKS) or do not cover all three HE families or are too mathematical. Therefore in this field, there is still a need for a comprehensively up-to-date survey which provides key concepts of the main encryption schemes in all three HE categories, together with their experimental performance comparison. The survey needs to be practical and show newly interested users how to build their own HE-based projects in popular HE libraries.

Our contribution

Our work aims to provide readers with fundamental principles of HEs without delving too deep into the mathematics. Furthermore, the thesis conducts a comprehensively theoretical and practical comparison of important HEs, covering all three HE categories: FHE, SWHE, and PHE. For different HE schemes in each family, we analyse their input parameters, together with their constraints, and then compare them together. This hands-on experience helps unprofessional practitioners distinguish libraries' properties and makes them easy to apply in building their own HE-based projects. In addition, we provide experimental results on performance evaluation of each HE scheme in most-used libraries such as SEAL [Lai17], PALISADE [Pol+22], HELib [HS14], and HEAAN [Che+21]. Although PHE schemes are now not available in mentioned open-source libraries, our own implementations of Paillier, El-Gamal, and RSA are used as partially homomorphic cryptosystems in the emulation. For each execution case, we also come up with assessments and results' explanations. Furthermore, in the last part, we deliver a concrete discussion on the security of aforementioned schemes against IND-CPA, IND-CCA, as well as integer factorization attacks on classical and quantum computers.

Authors	**Description**	**Scope**	
		Schemes	**Libraries**
C. Fontaine et al., 2007 [FG07]	Providing nonspecialists with a survey of HE techniques	-	-
T. Lepoint et al., 2014 [LN14]	Conducting a comparison of FV and YASHE schemes and explaining how to choose parameters to ensure correctness and security against lattice attacks	BGV YASHE	FLINT
V. Migliore et al., 2016 [MBF16]	Proposing a study of the current best solutions, providing a deep analysis of how to setup and size their parameters	BFV SHIELD	-
P. Martins et al., 2017 [MSM17]	Studying SWHE and FHE schemes supported by their performance and security from an engineering standpoint	BGV BFV Paillier El-Gamal	-
A. Acar et al., 2018 [Aca+18]	Providing a comprehensive survey of the main FHE, PHE and SWHE schemes, including the FHE implementations	RSA Paillier El-Gamal	SEAL HELib
B. Alaya et al., 2020 [ALM20]	Presenting different HE cryptosystems, joined with a final comparison between the adopted techniques	-	-
A. Kim et al., 2021 [KPZ21]	Revisiting BGV and BFV, together with proposing an improved variant of BFV	BGV BFV	PALISADE
C. Zaraket et al., 2021 [Zar+21]	Proposing SAVHO homomorphic scheme and its performance analysis in comparison with Pailler cryptosystem	Paillier SAVHO	SageMath
V. Sidorov et al., 2022 [SWN22]	Conducting an extensive study of homomorphic cryptosystems' performance for practical data processing	Paillier El-Gamal	SEAL PyAono HELib
S. J. Mohamme et al., 2022 [MT22]	Evaluating performance of RSA, El-Gamal, and Paillier homomorphic encryption algorithms	RSA El-Gamal Paillier	-

TABLE 2.1: Related works and their scopes.

Chapter 3

Background and Preliminaries

3.1 Libraries

HElib (**H**omomorphic-**E**ncryption **Lib**rary) [HS14] is the first open source library implementing HE. Being published in 2013, it focuses on effective use of BGV and CKKS schemes, together with ciphertext packing techniques and the Gentry-Halevi-Smart optimizations. Helib is still under development by Shai Halevi (IBM), Victor Shoup (NYU, IBM) and available on GitHub [HEl20]. In 2018, the authors implemented several algorithmic improvements, including Faster Homomorphic Linear Transformations[HS18], that made Helib 30–75 times faster than those previously built for typical parameters.

PALISADE [Pol+22] is a C++ open-source project, which provides efficient implementations of lattice cryptography building blocks. The library now supports varied HE schemes, such as: BGV, BFV, CKKS, FHEW, and TFHE. In addition, it also supports multi-party extensions of certain schemes and related cryptography primitives, namely digital signature schemes, proxy re-encryption, and program obfuscation. PALISADE is also available on GitHub [PAL20].

SEAL (**S**imple **E**ncrypted **A**rithmetic **L**ibrary) [Lai17] is also a HE library, developed by the Cryptography and Privacy Research Group at Microsoft. According to his author, Kim Laine, the first version of SEAL was released in 2015 with the specific goal of providing a well-engineered and documented HE library. SEAL was designed to use both by experts and by non-experts with little or no cryptographic background. The updated version of Microsoft SEAL, which can be found on GitHub [SEA20], has implemented various forms of HE schemes, including BGV, BFV, and CKKS. Besides, there is a SEAL version in Python, called SEAL - Python [SEA22]. This is a Python wrapper implementation of the SEAL library, using `pybind11` [pyb21].

HE scheme/Library	SEAL	PALISADE	HElib	HEAAN
BFV	✓	✓		
BGV	✓	✓	✓	
CKKS	✓	✓	✓	✓

Table 3.1: HE schemes in HE open-source libraries.

To have an extensive comparison for CKKS encryption, apart from these three mentioned libraries, we also measure its running time in HEAAN library [Che+21], developed in 2016 by its own authors. HEAAN (**H**omomorphic **E**ncryption for **A**rithmetic of **A**pproximate **N**umbers) is an open-source cross platform software library which implements the approximate HE scheme proposed by Cheon, Kim, Kim and Song

(CKKS). HEAAN executes only CKKS schemes with its complete properties. Following its owners, the library allows additions and multiplications to be performed by fixed point arithmetics and approximate operations between rational numbers. Table 3.1 illustrates the distribution of several encryption schemes in each library.

3.2 Homomorphic encryption schemes

In this part, we explain basic properties of HE, followed by a brief description of some notable PHE, SWHE, and FHE schemes, based on five main homomorphic operations: Key generation (`KeyGen`), Encryption (`Enc`), Decryption (`Dec`), Homomorphic addition (`Add`), and Homomorphic multiplication (`Mult`). The performance evaluation of mentioned schemes will be presented with greater detail in chapter 4, but first of all, we define a homomorphic encryption. In [Aca+18], an encryption scheme is called homomorphic over an operation "$\star$" (e.g., `Add`, `Mult`) if it supports the following equation:

$$E(m_1) \star E(m_2) = E(m_1 \star m_2), \forall m_1, m_2 \in M,$$

where E is the encryption algorithm and M is the set of all possible messages.

3.2.1 RSA

This HE was first introduced by Rivest et al. [RSA78]. The security of the cryptosystem relies on the practical hardness of factoring the product of two large prime numbers [Mon94], called the *factoring problem*. Given a security parameter λ, RSA is defined as follows:

- `KeyGen`(λ): First, two large prime numbers (p and q) are randomly chosen, then $N = pq$ and $\phi(N) = (p-1)(q-1)$ are computed. The secret large integer d is picked such that $\gcd(d, \phi(N)) = 1$. The last public component e is calculated by computing the multiplicative inverse of d (i.e., $ed \equiv 1 \mod \phi(N)$). Finally, set the public key $pk = (e, N)$, and the secret key $sk = (d, p, q)$.
- `Enc`($pk, m \in \mathbb{Z}_N$): The message m is an integer between 0 and $N-1$. The encryption of m is c, such that: $c = E(m) = m^e \pmod N$.
- `Dec`(sk, c): The message m can be recovered from the ciphertext c by: $m = c^d \pmod N$.
- `Mult`(c_1, c_2): $c_1 c_2 = E(m_1)E(m_2) = [m_1^e \pmod N][m_2^e \pmod N]$
  $= (m_1 m_2)^e \pmod N = E(m_1 m_2)$.

3.2.2 El-Gamal

The encryption system is a widely-used HE in public-key cryptography, proposed by T. ElGamal in 1985 [ElG85]. The advent of El-Gamal algorithm is based on the Diffie–Hellman key exchange, while its security strength is relied on the hardness of solving discrete logarithms.

- `KeyGen`(λ): Firstly a cyclic group G of order N and its generator $g \in \mathbb{Z}_N^*$ are generated. After randomly drawing an integer x from $\{1, \ldots, N-1\}$, $h = g^x$ is computed. The public key pk consists of (G, N, g, h), while $sk = x$ is kept secret.
- `Enc`($pk, m \in \mathbb{Z}_N$): A message m is encrypted by choosing an integer y randomly from $\{1, \ldots, N-1\}$, then computing $s = h^y$. the output of the encryption is a ciphertext $c = (c_1, c_2)$, where $c_1 = g^y$ and $c_2 = ms$.

- `Dec`(sk, c): To decrypt the ciphertext, firstly $s' = c_1^x$ needs to be calculated. Next, m is recovered by $m = c_2 s'^{-1}$.
- `Mult`(c_1, c_2):
$c_1c_2 = E(m_1)E(m_2) = (g^{x_1}, m_1h^{x_1}) \cdot (g^{x_2}, m_2h^{x_2}) = (g^{x_1+x_2}, m_1m_2h^{x_1+x_2}) = E(m_1m_2)$.

3.2.3 Paillier

The encryption of Paillier (1999) [Pai99] is an additively homomorphic cryptosystem, which is based on the composite residuosity problem and gathers many good properties.

- `KeyGen`(λ): Two primes numbers p, q of k bits are randomly generated such that $N = pq$ and $\rho = N^{-1} \pmod{\phi(N)}$, where $\phi(N) = (p-1)(q-1)$. One can publish $pk = N$ and $sk = \rho$.
- `Enc`$(pk, m \in \mathbb{Z}_N)$: To encrypt a message m, first an integer r from $\{1, \ldots, N-1\}$ is chosen randomly. The output is the ciphertext $c = (1 + mN)r^N \pmod{N^2}$.
- `Dec`(sk, c): To recover the message m, $r = c^\rho \pmod N$ is computed. Then $m = \frac{(cr^{-N} \pmod{N^2})-1}{N}$.
- `Add`(c_1, c_2): $c_1c_2 = E(m_1)E(m_2) = (1+m_1N)r_1^N(1+m_2N)r_2^N \pmod{N^2} = [1+(m_1+m_2)N + m_1m_2N^2](r_1r_2)^N \pmod{N^2} = [1+(m_1+m_2)N]r^N \pmod{N^2} = E(m_1+m_2)$.

3.2.4 BFV

In 2012, J. Fan and F. Vercauteren ported the scheme proposed by Brakerski [Bra12] from the learning-with-errors (LWE) setting to the Ring-LWE setting [FV12]. Using a simple modulus switching trick, BFV (so-called FV) simplifies the analysis of the bootstrapping step. The security of BFV-type cryptosystems is based on the LWE over rings (or RLWE) assumption [Reg09]. The RLWE(λ, q, χ) assumption states that it is very hard to distinguish two distributions $(a, b = a \cdot s + e)$ and (a, u), where a, s, and u are randomly selected from R_q and e is selected from an error distribution χ, referencing security parameter λ. This assumption has been proved hard over ideal lattices in [LPR10].

Let $R = \mathbb{Z}[x]/f(x)$ be a ring of polynomials in which the operations of BFV will be performed, where $f(x) = x^N + 1$ is a cyclotomic polynomial with N being a power of 2. The ring is used to define the RLWE problem with coefficients in $\mathbb{Z}_q$, denoted by $R_q = \mathbb{Z}_q[x]/f(x)$. Additionally, the message space is defined as being R_t for an integer $t > 1$.

- `KeyGen`(λ): For a B-bounded distribution χ over the ring R, a vector of secret key $sk = s$ is sampled $s \leftarrow \chi$. The public key is defined by: $pk = ([-(a \cdot s + e)]_q, a)$, where $e \leftarrow \chi$ and $a \leftarrow R_q$.
- `Enc`$(pk, m \in R_t)$: Given a plain message m, let $p_0 = pk[0]$, $p_1 = pk[1]$, and draw $u, e_1, e_2 \leftarrow \chi$, the ciphertext is: $c = ([p_0 \cdot u + e_1 + \Delta \cdot m]_q, [p_1 \cdot u + e_2])$, where $\Delta = \lfloor q/t \rfloor$.
- `Dec`(sk, c): Let $c = (c_0, c_1)$ be an encrypted message. The decryption returns m such as $m = \left[\left\lfloor \frac{t}{q}[c_0 + c_1 \cdot s]_q \right\rceil\right]_t$.

- `Add`(c_1, c_2): Let c_1, c_2 be two encrypted messages such that $c_1 = (c_{10}, c_{11})$ and $c_2 = (c_{20}, c_{21})$. The addition of two digits is $c = ([c_{10} + c_{20}]_q, [c_{11} + c_{21}]_q)$.
- `Mult`(c_1, c_2): By multplying two ciphertexts $c_1(s)$ and $c_2(s)$, the result is $c_1(s) \cdot c_2(s) = c'_0 + c'_1 \cdot s + c'_2 \cdot s^2$. One encountered problem is that resulting ciphertext has degree 2 and must be reduced to a degree 1 [FV12]. This process is called *relinearization*, which is similar to modulus switching. To start, a relinearization key rlk is generated by choosing an integer p and sampling a new $a \leftarrow R_{pq}$ and $e \leftarrow \chi'(\chi' \neq \chi)$ satisfying $rlk = ([-(a \cdot s + e) + p \cdot s^2]_{pq}, a)$. The output is relinearizated 1-degree ciphertext: $([c_0 + c_{2,0}]_q, [c_1 + c_{2,1}]_q)$, where:

$$c_0 = \left[\left\lfloor \frac{t(c_{10} \cdot c_{20})}{q} \right\rceil\right]_q$$

$$c_1 = \left[\left\lfloor \frac{t(c_{10} \cdot c_{21} + c_{11} \cdot c_{20})}{q} \right\rceil\right]_q$$

$$c_2 = \left[\left\lfloor \frac{t(c_{11} \cdot c_{21})}{q} \right\rceil\right]_q$$

$$(c_{2,0}, c_{2,1}) = \left(\left[\left\lfloor \frac{c_2 \cdot rlk[0]}{p} \right\rceil\right]_q, \left[\left\lfloor \frac{c_2 \cdot rlk[1]}{p} \right\rceil\right]_q\right)$$

3.2.5 BGV

BGV encryption was invented in 2011 by Brakerski, Gentry, and Vaikuntanathan [BGV14]. BGV is a FHE that works for both an LWE and an RLWE. The hardness of the scheme is also based on RLWE problem [LPR10]. To keep the ciphertext error within a given bound, they used the technique of modulus switching as introduced in [BV14]. This modulo reduction maps a ciphertext c defined in a ring R_q, to a ring R_p, where $p < q$ [RLF19]. This keeps the error contained within the ciphertext at the same level. In original BGV, public key and switch keys are matrices [YG20]. Given a security parameter λ and level L, corresponding to the maximum number of multiplications that can be executed. First step is to generate L large primes $q_0, \dots q_{L-1}$ satisfying $q_0 < \dots < q_{L-1}$ [YG20]. p is also chosen as a plaintext modulus.

- `KeyGen`(λ, χ, L): A vector s is selected randomly as a sk. Then $b = -(a \cdot s + p \cdot e) \pmod{q_{L-1}}$ is computed, where $a \leftarrow R_{q_{L-1}}$, $e \leftarrow \chi$. The public key is (a, b). Next, the switch keys $(a_0, b_0, t_0, 0), \dots, (a_{L-1}, b_{L-1}, t_{L-1}, L-1)$ will be computed, where $b_i = -(a_i \cdot s + p \cdot e_i - t_i \cdot s^2) \pmod{t_i \cdot q_i}$, $a_i \leftarrow R_{q_i}$, $e_i \leftarrow \chi$, and t_i is an integer.
- `Enc`$(pk, m \in \mathbb{Z}_p)$: A plaintext m can be encrypted by $E(m) = (c_0, c_1) = ((b \cdot v + p \cdot e_0 + m) \pmod{q_{L-1}}, a \cdot v + p \cdot e_1 \pmod{q_{L-1}})$, where each element of vector v, $v_i \in \{0, 1, -1\}$ and $e_0, e_1 \leftarrow \chi$. We have $c = (c_0, c_1, L-1)$ is the initial ciphertext.
- `Dec`(sk, c): A ciphertext $c = (c_0, c_1, i)$, $i = [0, L-1]$ can be decrypted to find its plaintext m by $m = c_0 + c_1 \cdot s \pmod{q_i} \pmod{p}$.
- `Add`(c_1, c_2): The addition of two ciphertexts of a same level $c_1 = (c_{10}, c_{11}, i)$ and $c_2 = (c_{20}, c_{21}, i)$ is computed by $c = ((c_{10} + c_{20}) \pmod{q_i}, (c_{11} + c_{21}) \pmod{q_i}))$.
- `Mult`(c_1, c_2): Similarly to BFV method, the first step is to compute a degree-2 ciphertext, where we denote $c_1(s) \cdot c_2(s) = c'_0 + c'_1 \cdot s + c'_2 \cdot s^2$. The relinearization procedure results a compressed ciphertext with degree 1: $c^* = (c^*_0, c^*_1)$, where $c^*_0 = t_i c'_0 + b_i \cdot c'_2 \pmod{t_i q_i}$ and $c^*_1 = t_i c'_1 + a_i \cdot c'_2 \pmod{t_i q_i}$, with the switch

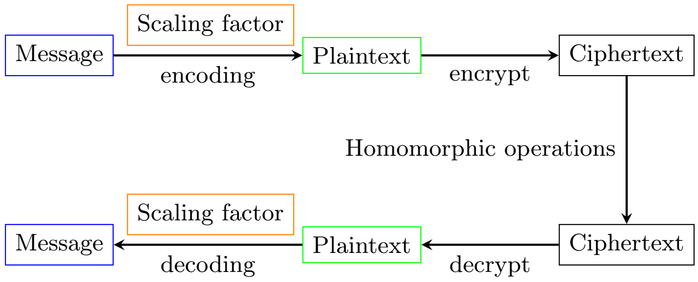

FIGURE 3.1: Algorithms in CKKS.

key (a_i, b_i, t_i, i). The ciphertext c^* will be mapped to $c \in R_{q_{i-1}}$ as the output by *SwitchModulus* method.

- `SwitchModulus`$(c = (c_0, c_1, i))$: Supposing to have two modulus q_i and q_j where $i > j$, and a ciphertext c in ring R_{q_i}, we calculate modulo inverse element $r_j = \frac{q_j}{q_i}$ in q_j. The new ciphertext in ring R_{q_j} is defined by $\bar{c} = (\bar{c}_0, \bar{c}_1, j) = (c_0 r_j \pmod{q_j}, c_1 r_j \pmod{q_j}, j)$.

3.2.6 CKKS

As mentioned in the previous section, CKKS, a HE for approximate arithmetic, was introduced in 2016 in [Che+17]. What makes CKKS draw attention to many researchers is that it allows to perform approximate additions and multiplications of ciphertexts, where its plaintexts can be vectors of real and complex values. This has been done by *encoding* and *decoding* method, where the inputs are converted from $C^{N/2} \times \mathbb{R}$ to $R = \mathbb{Z}[x]/(x^N + 1)$ and vice versa [Che+17]. In this step, we need to use a rounding technique, which might destroy some significant numbers. Thus, if we had an initial vector of real or complex values z, roughly speaking it will be multiplied by a scale $\Delta > 0$ during encoding and then divided by Δ during decoding to keep a precision of $\frac{1}{\Delta}$. Figure 3.1 describes all algorithms in CKKS scheme [Yon19].

As well as many other HE schemes, the foundation of CKKS is also the RLWE problem. Similarly to previously presented schemes, in this part, we simply describe the five main algorithms of CKKS. To start, it begins with a integer $p > 0$, number of multiplication L, and modulus q_0. For $0 < l \leq L$, we define $q_l = p^l q_0$.

- `KeyGen`(λ, q_L): First, a vector s is sampled from a set of signed binary vectors in $\{0, 1, -1\}^N$ whose Hamming weight is exactly an integer h. Next, $a \leftarrow R_{q_L}$, and $e \leftarrow \chi$. We set the secret key $sk = (1, s)$, $pk = (b, a) \in R^2_{q_L}$ with $b = -as + e \pmod{q_L}$. We choose an integer P, set $a' \leftarrow R_{P \cdot q_L}$, $e' \leftarrow \chi$, and $evk = (b', a') \in R^2_{P \cdot q_L}$ with $b' = -a's + e' + Ps^2 \pmod{P \cdot q_L}$.
- `Enc`(pk, m): Given a distribution $ZO(\rho)$ draws each entry in the vector from $\{0, 1, -1\}^N$, with probability $\rho/2$ for each of -1 and $+1$, and probability being zero $1 - \rho$. To encrypt a polynomial m, we sample polynomials $v \leftarrow ZO(0.5)$, $e_0, e_1 \leftarrow \chi$, then output the ciphertext $c = v \cdot pk + (m + e_0, e_1) \pmod{q_L}$.
- `Dec`(sk, c): For a ciphertext $c = (b, a) \in R^2_{q_l}$, the approximate result m' of the plaintext m can be recovered by $m' = m + e = b + a \cdot s \pmod{q_l}$.
- `Add`(c_1, c_2): For $c_1, c_2 \in R^2_{q_l}$, its addition is $c = c_1 + c_2 \pmod{q_l}$.
- `Mult`(evk, c_1, c_2): Similarly to introduced HEs, the multiplication of CKKS also accompanies a relinearlization step. For $c_1 = (b_1, a_1)$, $c_2 = (b_2, a_2) \in R^2_{q_l}$, let

$(c'_0, c'_1, c'_2) = (b_1 b_2, a_1 b_2 + a_2 b_1, a_1 a_2) \pmod{q_l}$. After being relinearlizated, it outputs a degree-1 ciphertext $c = (c'_0, c'_1) + \lfloor P^{-1} \cdot c'_2 \cdot evk \rceil \pmod{q_l}$.

One problem produced is that underlying value contained in the plaintext and ciphertext is $\Delta \cdot z$ as mentioned above. So after multiplying two ciphertexts c_1, c_2, the result holds $z_1 \cdot z_2 \cdot \Delta^2$. By doing many multiplications, the resulting ciphertext will have grown exponentially. To reduce its size, *Rescale* $RS_{l \to l'}$ is introduced with its goal being to actually keep the scale constant, and also reduce the noise present in the ciphertext.

- $RS_{l \to l'}(c)$: For a ciphertext $c \in R_{q_l}^2$ at level $l > l'$, we output $c' = \left\lfloor \frac{q_{l'}}{q_l} c \right\rceil \in \pmod{q_{l'}}$.

The schemes presented above are six homomorphic cryptosystems within the scope of the project, which are also the most widely used in cryptography applications and the basis for other schemes. In the following chapter, we will elaborate their implementation methods and results of evaluation analysis based on practical experiments.

Chapter 4

Implementation and results

4.1 Notations

The main focus of this article is to compare each available scheme's performance in different libraries. For this reason, in each library, we built our own "simple" project as a regular end-user. Each project is corresponding to one scheme, which includes five main homomorphic operations: `KeyGen`, `Enc`, `Dec`, `Add`, and `Mult`. The execution time needed to perform each operation will be recorded and then compared to each other. Every program collecting the performance metrics is carried out on an average commodity computer equipped with with an Intel(R) Core(TM) i7-10700 CPU running at 2.90GHz under Ubuntu 20.04. In the results presented in the next section, the following notations are used:

Symbol	Description
p	**The plaintext modulus** of BGV, BFV schemes
Q	**The maximum ciphertext modulus**, the initial ciphertext modulus after encryption
m	**The cyclotomic order** of the ring $\mathbb{R}$
N	**The degree** of the ring $\mathbb{R}$ ($N = \phi(m)$)
n	**The number of slots** or messages encoded in one ciphertext
L	**Multiplication depth**, the number of multiplications can be executed
Δ	**Scaling factor** in CKKS scheme, multiplied to the floating-point number of message to convert to integer number.

To ensure the consistency in test results, every experiment is executed according to the strategy below:

- The time unit is microseconds (μs);
- Each operation was executed in 1000 iterations and the time presented is its average;
- The parameters are chosen to ensure the 128-bit encryption security level;
- The time measured of encryption operation includes the execution time of random values for message inputs, together with encoding and decoding timings for batching;
- Bootstrapping is not applied.

Depending on different HE schemes' properties, chosen plaintext will be differed. BFV and BGV schemes allow modular arithmetic on encrypted integers, while CKKS supports homomorphic operations on real or complex ones. Within the scope of the project, plaintext batching technique is applied for all evaluated FHE and SWHE schemes. The first element of the batch is drawn randomly from a uniform distribution over the same range p, and the remaining elements are set to be 0. The diversity in input setting for each scheme will be explained with greater detail in the following subsections.

4.2 Fully homomorphic encryption

4.2.1 BFV

Although the scheme is available in both SEAL and PALISADE as mentioned in Table 3.1, they also have their differences in the implementation, indicated in Table 4.1. PALISADE allows users to change the parameters p, N, L as inputs, whereas SEAL-Python keeps L unchangeable from the user side. To accelerate the batching technique, the two libraries require that the chosen plaintext modulus p needs to be a prime number and congruent to 1 (mod $2n$). This is the condition to operate on n packed integers in a SIMD (*Single Instruction, Multiple Data*) manner [Lai17]. In order to assess the relative practical efficiency of two libraries for BFV encryption, different implementations are done with the same input parameters and working environment given in Table 4.2.

		SEAL-Python	**PALISADE**
Languages		Python	C++
Parameters	p	changeable	changeable
	N	changeable	changeable
	L	not changeable	changeable
	Q	not changeable	not changeable
Batching		$n = N$	$n = N$
Condition		$p = 1 \pmod{2n}$, p is a prime number	

TABLE 4.1: Differences in libraries' setups.

p	$log_2 Q$	required N
1032193	109	4096
1032193	218	8192
786433	438	16384
786433	881	32768

(A) SEAL

p	$(log_2 Q, L)$	required N
1032193	(120,1)	4096
1032193	(180, 2), (180, 3)	8192
786433	(240,4), (300,5), (300,6), (360,7), (360,8), (420,9)	16384
786433	(480,10), . . . , (780,19), (840,20), (840,21)	32768

(B) PALISADE

TABLE 4.2: BFV's input parameters in PALISADE and SEAL.

Unlike SEAL, PALISADE can calculate required N and Q based on chosen L and p to ensure a security level of 128 bits. In contrast, SEAL sets 128-bit encryption security level as default and allows users to enter N. SEAL then displays satisfied Q and p with the entered inputs. Table 4.2 contains many options for PALISADE's inputs with the same p and N in order to have 128-bit security. However, to have fair comparison between these two, the value of Q in PALISADE is chosen to be close to the one in SEAL. In Table 4.3, we provide timings for five main cryptographic functions, using the parameters recommended in Table 4.2.

HE parameters		KeyGen	Enc	Dec	Add	Mult
N	log_2Q					
4096	109	1028.119	1263.528	276.045	1.298	3274.257
8192	218	3003.509	3269.548	1179.682	144.531	11663.16
16384	438	10260.45	11378.441	5434.016	415.662	54918.967
32768	881	40251.496	41297.274	17442.857	1536.587	246427.201

(A) SEAL

HE parameters		KeyGen	Enc	Dec	Add	Mult
N	log_2Q					
4096	120	1137.556	1160.459	283.99	0.237	4296.438
8192	180	3170.82	2881.717	921.646	187.703	13585.75
16384	420	13507.743	11288.5535	3298.9775	1086.105	76565.506
32768	840	55941.007	45587.262	17171.713	7046.362	427795.343

(B) PALISADE

TABLE 4.3: Horizontal comparison of BFV's execution time.

After examining these tables, it is clear that the ciphertext dimension N has a significant effect on BFV's performance. In most cases, the running times of decryption and addition are less than the others. In general, when N increases, the execution times of all operations are increased, especially multiplication, which approximately grows up 4 times compared to the previous N in both two libraries. However, in particular, the mean multiplication execution time of SEAL is less than that of PALISADE. One explanation for this is that the latter always counts the relinearization procedure whenever doing multiplication (`EvalMult` function), while in the former, it is separately computed. Within the scope of our experiments, the decryption is executed only on a fresh ciphertext without doing multiplication before. Therefore, it is not necessary to do relinearization step. That is why the timing in SEAL does not involve relinearization.

In Figure 4.1, we depict experimental results in vertical comparison, where timings are illustrated based on each operation. It is obvious that the mean execution times of all cryptographic functions in two libraries are close to each other, but SEAL is still performing better. While the rest are almost similar, the biggest variance is displayed in multiplication time, where $N = 32768$, SEAL is approximately 2 times faster than PALISADE.

4.2.2 BGV

Unlike BFV, all three libraries SEAL, PALISADE, and HElib have implemented BGV. Generally, in doing the experiments, the encryption parameters of BGV, namely p, Q, N, and security level, are kept unchanged, compared to BFV. The number of slots in one batch is $n = N = \phi(m)$, except for the last case of Helib where $N = 32768$ and $n = 8192$ as indicated in Table 4.4. This number is impacted by several

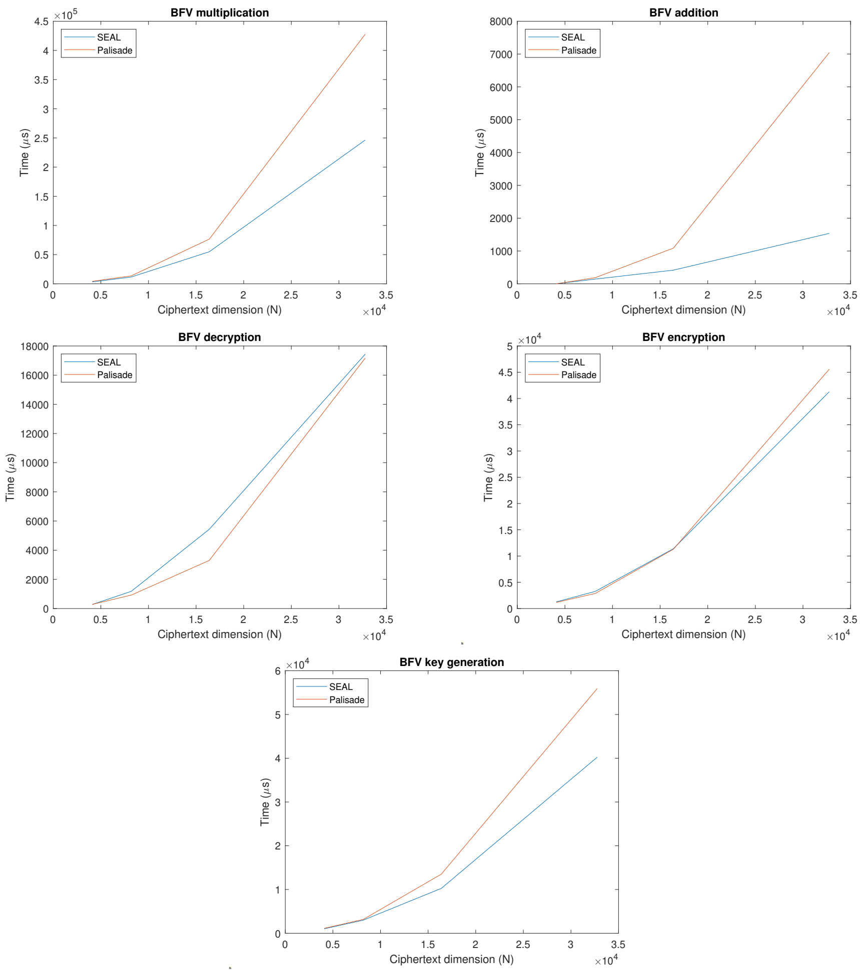

FIGURE 4.1: Vertical comparison of BFV's execution time.

parameters, including the maximum supported computation depth of the circuit (L) [HS14]. As L is varied to allow more computation, it also affects the cost of the computation. Additionally, in practical implementation, some technical definitions have been introduced in HE libraries, *noise budget* is one of them. According to A. Kim [Lai17], noise budget (invariant) is defined as the total amount of noise we have left until decryption will fail. To be more precise, the BGV implementation for each library is specified as follows:

PALISADE: In the library, noise budget is managed by `ModReduceInPlace` function, a method for reducing modulus of ciphertext and the private key used for encryption [Pol+22]. As explained above, in our scope of evaluation, this function will not be included. For BGV multiplication, the BFV operation of `EvalMult` is reused, so key-switching or relinearlization is already added. Besides, other properties of BGV implementation are remained the same as BFV's, such as the solution to calculate required N, Q, as well as the condition of inputs as mentioned in Table 4.1.

HElib: Helib allows to calculate its security level based on p, m, and `bits` (the number of bits of the modulus chain). When `bits` increase, its execution time is also raised up. Thus, in the comparison with other libraries as demonstrated on Table 4.4, we choose these variables such that the security level is close to or at least 128 bits.

Table 4.4 shows that the performance of HElib can be considered as good as the other two libraries if the timings of key generation and decryption were not such

HE parameters		KeyGen	Enc	Dec	Add	Mult
N	log_2Q					
4096	109	2424.838	1091.586	259.842	42.541	1508.681
8192	218	11426.94	3137.433	992.5	79.952	6673.09
16384	438	70869.416	11179.579	3791.998	292.17	35650.547
32768	881	433638.89	41716.827	18156.642	866.2635	215414.681

(A) SEAL

HE parameters		KeyGen	Enc	Dec	Add	Mult
N	log_2Q					
4096	96	3023.297	1145.76	368.375	42.116	570.952
8192	144	10981.757	3043.417	1007.424	57.322	2396.688
16384	240	51708.9	8902.513	3546.961	289.751	13642.014
32768	480	376273.9767	34662.558	20727.674	3313.547	116248.311

(B) PALISADE

HE parameters		KeyGen	Enc	Dec	Add	Mult
N	log_2Q					
4096	100	168300.764	2257.432	138092.51	32.064	2347.865
8192	100	470367.195	4533.877	549616.633	480.44	4492.487
16384	100	1348552.91	9917.878	2265994	289.706	10778.79
32768	100	1967110.87	14080.4445	2340201.2	209.039	17477.661

(C) HElib

TABLE 4.4: Horizontal comparison of BGV's execution time.

N	`addSome1DMatrices`	KeyGen	Enc	Dec	Add	Mult
4096	No	4882.369	2229.431	138098.495	120.27	2093.809
	Yes	168300.764	2257.432	138092.51	32.064	2347.865
8192	No	9903.988	4637.854	548176.341	389.5065	5463.149
	Yes	470367.195	4533.877	549616.633	480.44	4492.487
16384	No	20361.114	9418.236	2213333.56	577.842	11731.862
	Yes	1348552.91	9917.878	2265994	289.706	10778.79
32768	No	39825.16	13217.164	2373608.58	1039.197	21059.138
	Yes	1967110.87	14080.4445	2340201.2	209.039	17477.661

TABLE 4.5: BGV in HElib with different inputs.

slow. To explain this, we need to examine the execution of key-switching matrices `addSome1DMatrices` in `KeyGen` process. Table 4.5 displays the differentiation in running time of computing or not the `addSome1DMatrices` function. Without adding this procedure, key generation has been much less time-consuming. For instance, in case $N = 32768$, it took almost 2 seconds to generate its key pair with `addSome1DMatrices`, whereas this process costed only 40 milliseconds approximately without it.

In contrast, key-switching matrices have not been mentioned in SEAL and PALISADE. Instead, PALISADE calculates required N and Q as illustrated in Table 4.6.

The different pairs of L and Q in each line have the same level of security. Hence, in the horizontal comparison of Table 4.4, we selected PALISADE results with lower (L, Q) to compare with others. On the other hand, Table 4.7 contains the timing results when implementing the lowest and highest pairs of (L, Q) in each particular case of N value. Based on its behaviors, (L, Q) shows an impressing effect on PALISADE's execution time, especially on `KeyGen` procedure. For instance, at the same level $N =$

p	(log_2Q, L)	required N
1032193	(96,1)	4096
1032193	(144, 2), (192, 3)	8192
786433	(240,4), (288,5), (336,6), (384,7), (432,8)	16384
786433	(480,9), (528, 10), (576,11),. . . (768,15), (816,16), (17,864)	32768

TABLE 4.6: BGV inputs for 128-bit security level in PALISADE.

32768, $L = 17$ took more than 1 second to generate key pair, whereas 0.3 seconds is its cost when $L = 9$. Last but not least, Figure 4.2 exposes the visibly vertical comparison of the two based on timings of each operation.

N	L	log_2Q	KeyGen	Enc	Dec	Add	Mult
8192	2	144	10981.757	3043.417	1007.424	57.322	2396.688
	3	192	18055.952	3662.841667	1499.093333	77.27016667	3906.558667
16384	4	240	51708.9	8902.513	3546.961	289.751	13642.014
	8	432	151107.142	13704.344	9176.451	1203.262	44985.406
32768	9	480	376273.9767	34662.558	20727.674	3313.547	116248.311
	17	864	1138783.22	58661.984	52499.905	3221.175	362749.472

TABLE 4.7: BGV in PALISADE with different inputs.

A deep analysis of the Figure 4.2 and Table 4.4 shows that the SEAL and PALISADE are performing much better than the HElib for `KeyGen` and `Dec` operations. In contrast, Helib running time is the best in multiplication and encryption. On the other hand, PALISADE and SEAL have equally good performance in all operations. Although there is dissimilarity between them in multiplication and addition, since the actual time counted in μs, it is not really a great distance.

4.3 Somewhat homomorphic encryption

The CKKS scheme is called *leveled homomorphic encryption*, an "extended" form of SWHE. In contrast to BFV and BGV encryption, where exact values are necessary, CKKS allows both additions and multiplications on encrypted complex numbers, but yields only approximate results [Lai17]. According to A. Kim [SEA20], one should take advantage of CKKS encryption in applications such as summing up encrypted real numbers, evaluating machine learning models on encrypted data, or computing distances of encrypted locations. As a result, CKKS scheme has been implemented in four HE libraries as communicated in Table 3.1. To perform experiments with CKKS, in addition to the default setting mentioned above, the input parameters are the same for all libraries, where:

- Scaling factor $\Delta = 2^{40}$;
- For batching technique, $n = N/2$.

In this encryption, there is no condition of plaintexts. Based on its properties, we chose inputs as real numbers. The method to draw packed messages keeps unchanged as discussed in section 4.1.

PALISADE: The ring dimension of the HE scheme is chosen following the security standards. Hence, to meet a requirement of 128-bit security level, the minimum value of N is 8192. In Table 4.8, we presents the detail of input parameters.

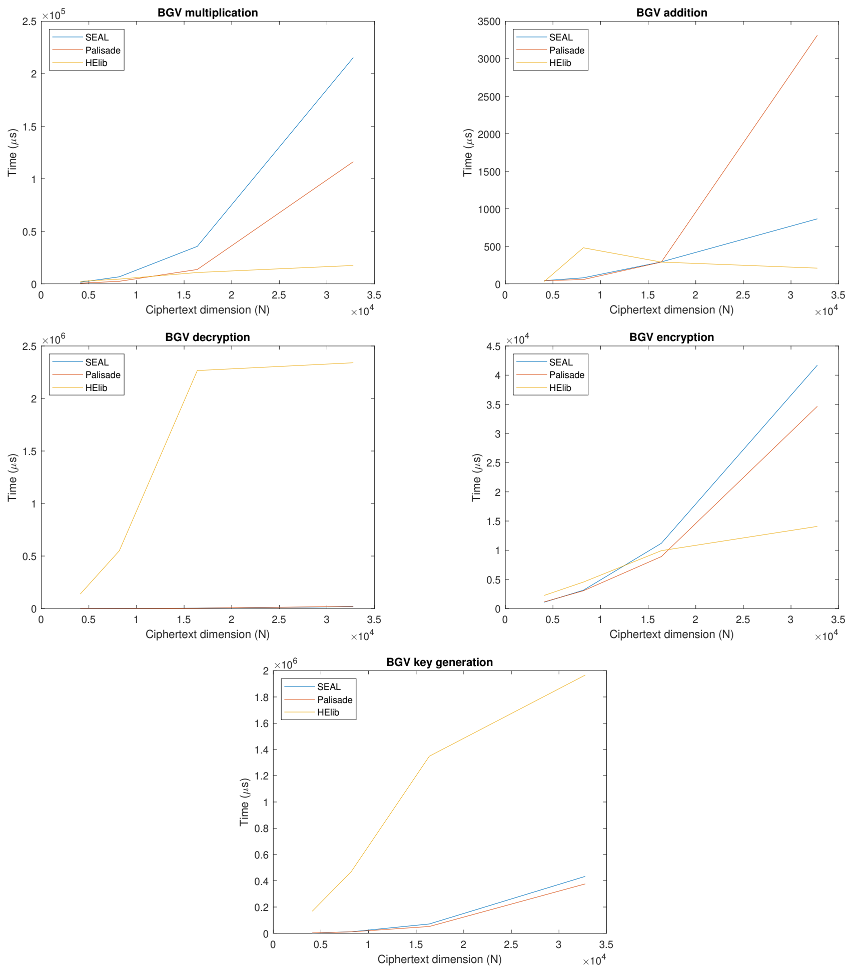

FIGURE 4.2: Vertical comparison of BGV's execution time.

SEAL: Like other CKKS implementation, SEAL does not use the plaintext-modulus parameter p. Moreover, instead of providing a ciphertext modulus Q, users working with CKKS must provide a modulus chain of prime sizes (e.g., $q = [60, 40, 40, 60]$) [SEA20]. The number of moduli is equal to the number of iterations/multiplications. Additionally, the $log_2 Q$ bit as shown in Table 4.2 is kept unchanged, but now it is corresponding to the maximal sum of these primes, called `CoeffModulus`.

Before going to the evaluation part of different libraries' performance, Table 4.9 illustrates how SEAL behaves sensitively with ciphertext modulus and its modulus composition for each value of N. Although addition and decryption time are not changed significantly, the calculation time is climbed up more than 2 times in the three remaining operations.

HElib: One of the most advantages of HElib is its transformation of complex mathematical calculations in order to be easier and more understandable for non-expert practitioners. For example, to add two ciphertexts `cipher_a` and `cipher_b`, HElib supports to simply declare a new one as a sum of the two: `Ctxt cipher_add = cipher_a; cipher_add += cipher_b`. There is no need to specify technical steps such as relinearization or rescaling as others. Being different from other libraries, HElib allows users to calculate encryption security level based on input parameters. Table 4.10 contains the experimental results with the encryption security `sec_level` being the closest to 128-bit level, while still preserving HElib's usage recommendation.

(log_2Q, L)	required N
(101,1)	8192
(140,2), (181,3), (221,4), (261,5), (301,6)	16384
(341,7), (381,8)	32768

TABLE 4.8: CKKS input parameters in PALISADE.

N	log_2Q	Modulus chain	KeyGen	Enc	Dec	Add	Mult
8192	160	60, 40, 60	2179.13	3206.887	50.54	128.72	178.348
	200	60, 40, 40, 60	2507.607	3910.8775	109.5735	271.207	452.792
16384	200	60, 40, 40, 60	5034.227	8548.111	221.213	317.991	774.644
	432	60, [39]*8, 60	11959.254	18847.575	721.076	1777.956	2077.872
32768	200	60,40,40,60	10215.043	18231.808	781.699	529.59975	1414.666
	881	[55]*15,56	39749.74	66061.559	2589.904	2221.2535	4741.014

TABLE 4.9: SEAL's comparison for different modulus composition.

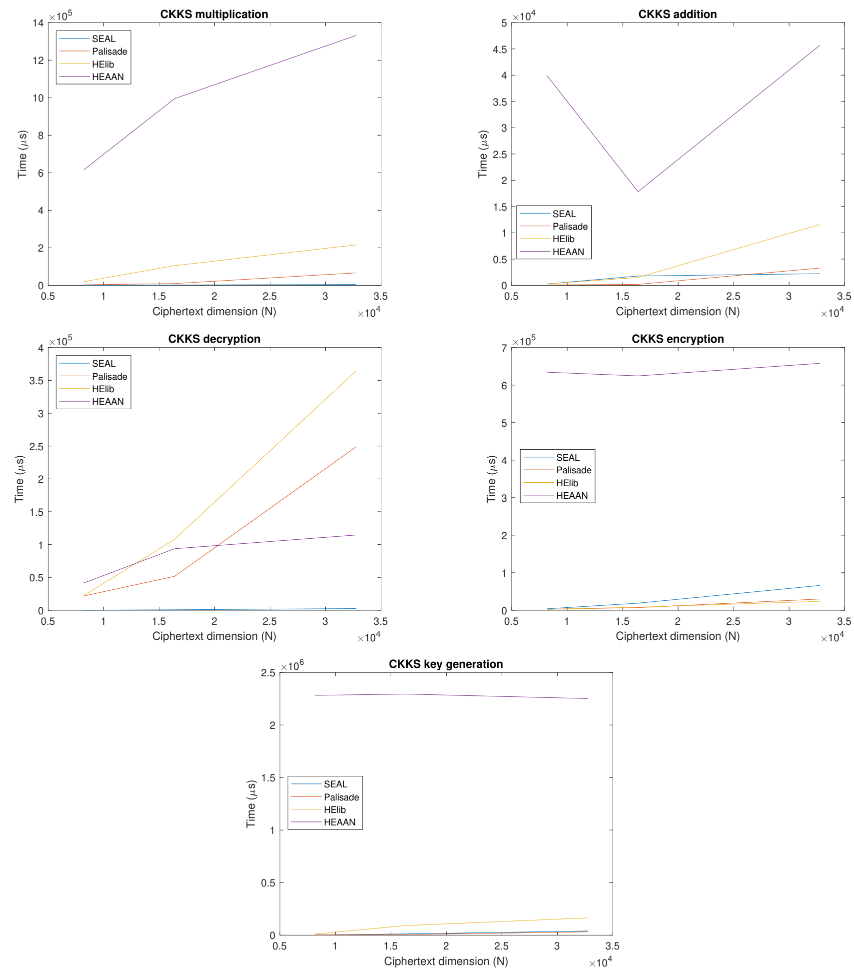

FIGURE 4.3: Vertical comparison of CKKS's execution time.

HEAAN: The last library was developed by its own authors. HEAAN takes advantage of fully built-in algorithms, where it is able to deal with complex numbers. An input message in HEAAN can consist of n complex numbers, where $n \leq N/2$.

By analyzing different results displayed in Figure 4.3, one can see that overall performance of HEAAN and Helib are quite slower than SEAL and PALISADE. In overall, SEAL owns the best performance, while HEAAN is much more time-consuming compared to the others. Considering PALISADE's presentation in both Table 4.10 and Figure 4.3, it is obvious that its most time-consuming procedure is decryption and multiplication. One reason needed to bring up is that, the relinearization step is always included in multiplication function `EvalMult`. Moreover, for decryption process (`cc->Decrypt(keys.secretKey, cMul, &result)`), calculated time of rescaling algorithm is also taken into account. In contrast, SEAL does not include relinearization and re-scaling schemes in multiplication and decryption respectively. Instead, it can be done by coding separately with `evaluator.relinearize_inplace` and `evaluator.rescale_to_next_inplace` functions.

HE parameters		KeyGen	Enc	Dec	Add	Mult
N	log_2Q					
8192	200	2507.607	3910.8775	109.5735	271.207	452.792
16384	432	11959.254	18847.575	721.076	1777.956	2077.872
32768	881	39749.74	66061.559	2589.904	2221.2535	4741.014

(A) SEAL

HE parameters		KeyGen	Enc	Dec	Add	Mult
N	log_2Q					
8192	102	2305.699	2652.975	21650.503	81.117	3129.505
16384	141	6542.385	7093.977	51639.085	194.05	9584.286
32768	342	31630.72	29936.449	248985.192	3291.783	66603.916

(B) PALISADE

HE parameters		KeyGen	Enc	Dec	Add	Mult
N	(log_2Q,sec_level)					
8192	(119,157.866)	11008.069	2659.019	22065.082	272.865	19712.186
16384	(358,129.741)	91768.896	8252.838	107935.827	1502.701	104850.697
32768	(558,128.851)	164575.383	23730.201	364743.317	11576.171	215878.991

(C) HElib

HE parameters		KeyGen	Enc	Dec	Add	Mult
N	log_2Q					
8192	119	2282102.44	634268.04	41491.42	39877.65	614878.85
16384	358	2294477.86	624440.22	93658.41	17826.4	994892.6
32768	558	2251482.12	657943.99	114587.91	45690.59	1332368.41

(D) HEAAN

TABLE 4.10: Horizontal comparison of CKKS's execution time.

4.4 Partially homomorphic encryption

This part presents our own implementations of partially homomorphic cryptosystems, including Paillier (additive), El-Gamal (multiplicative), and RSA (multiplicative). The source code is available at `github` [Doa22]. Table 4.11 and Figure 4.4 illustrate horizontal and vertical comparison results respectively. According to PHE's properties as introduced in chapter 1, one PHE scheme can possess four following operations: Key generation, encryption, decryption, and addition/multiplication. Unlike FHE and SWHE, here the inputs are identified as p (plaintext modulus) and log_2N (the number of bits of N), where N is one factor in public (encryption) keys. For each cryptosystem, we measure the execution time when selecting pairs of $(log_2 N, p)$ in the similar manner of selecting $(log_2 Q, p)$ in FHE. In addition, we execute the second situation where p is in $\mathbb{Z}_N$ and the bits of N are so large such that they can reach 128-bit security level as stated in [Hea+13]. As same as the previous implementations, the time unit is microseconds; each operation was executed in 1000 iterations and the timings presented are its average. The implementations are set up following their original papers: Paillier [Pai99], El-Gamal[ElG85], and RSA [RSA78]. Table 4.11 and Figure 4.4 below demonstrate the experimental results.

HE parameters		KeyGen	Enc	Dec	Mult
p	$log_2 N$				
1032193	109	1484.324	1.062	5.399	2.237
1032193	218	1931.446	1.669	7.803	0.396
786433	438	3490.179	2.496	37.265	0.853
786433	881	8366.837	6.865	205.38825	2.178
$\mathbb{Z}_N$	3072	180255.34	61.599	6327.537	2.934
$\mathbb{Z}_N$	4096	433348.8	88.372	14327.857	9.792

(A) RSA encryption

HE parameters		KeyGen	Enc	Dec	Mult
p	$log_2 N$				
1032193	109	31063.45	4.486	4.336	3.151
1032193	218	135618.53	16.748	15.476	9.74
786433	438	773368.775	66.794	32.56	18.388
786433	881	5354554.333	403.448	203.718	8.651
$\mathbb{Z}_N$	3072	>15 minutes			
$\mathbb{Z}_N$	4096	>15 minutes			

(B) El-Gamal encryption

HE parameters		KeyGen	Enc	Dec	Add
p	$log_2 N$				
1032193	109	1072.014	265.509	6.738	4.255
1032193	218	1537.688	279.664	22.872	4.053
786433	438	3081.14	367.893	141.012	6.08
786433	881	7903.175	1013.957	950.735	10.868
$\mathbb{Z}_N$	3072	183774.01	20237.659	26364.306	232.136
$\mathbb{Z}_N$	4096	4297885.6	42868.889	55843.731	212.978

(C) Paillier encryption

TABLE 4.11: Horizontal comparison of PHE's execution time.

4.4.1 Paillier encryption

In Paillier cryptosystem, although the input is N, the cipher space or ciphertext modulus is N^2 (see chapter 3). In spite of that, generally the algorithm performs all four operations very well as shown in Table 4.11. In the second situation, when both p and N increase, there is no much difference in time execution of `Add`. However, for three others, they are both climbed up. Particularly, when N is 4096 bits, the average time for one `KeyGen` is almost 4.3 seconds.

4.4.2 El-Gamal encryption

Table 4.11 indicates that all three operations of encryption, decryption, and multiplication in El-Gamal method have a better performance compared to Paillier. Apart from that, `KeyGen` appears to be a very time-consuming procedure. The cryptosystem needs more than 5 seconds to generate key pairs if $log_2 N = 881$, not mention to say that it needs more than 15 minutes when $log_2 N = 3072$ or more. Regarding to this problem, its author Taher ElGamal explained that in any of the cryptographic systems based on discrete logarithms like El-Gamal, N must be chosen such that $N - 1$ has at least one large prime factor [ElG85]. If $N - 1$ has only small prime factors, computing discrete logarithms would be easy [PH78]. Hence, our implementation is set up such that this condition is satisfied. N is considered as a safe prime if $(N - 1)/2$ is also a prime.

4.4.3 RSA encryption

It is clearly seen in Figure 4.4 that RSA has represented the best performance among three PHE schemes, even in case of very large ciphertext space. As its authors stated in [RSA78], the secret key d in RSA is very easy to choose, which is relatively prime to $\phi(N)$, where $N = pq$. To be more specific, any prime number greater than $\max(p, q)$ will do. This is one of the reasons why RSA does not take much time to generate keys like El-Gamal encryption and why it is commonly used in practice.

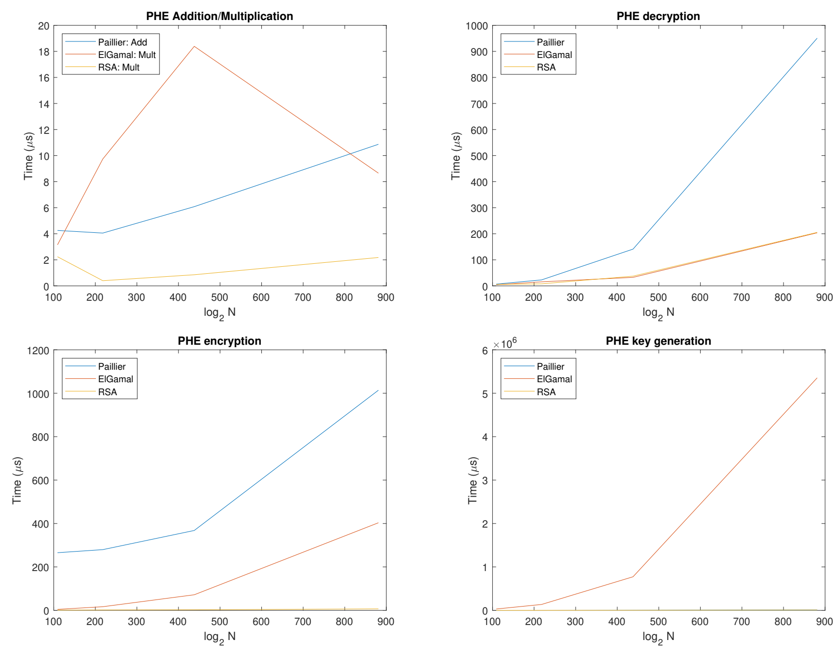

FIGURE 4.4: Vertical comparison of PHE's execution time.

In Figure 4.4, the first graph on the top-left side displays the running time of Addition (`Add`) in Paillier and Multiplication (`Mult`) for the remaining two cryptosystems. Although the difference among them is demonstrated visibly, it is still considered as marginally small for the time unit is in μs.

In this chapter, we made an efficient performance comparison of six notable HE schemes, covering all three homomorphic encryption categories: Partially HE, Somewhat HE, and Fully HE. The results clearly suggest that partially homomorphic cryptosystems are significantly faster than the others at addition and multiplication operations. RSA possesses the fastest key-generation procedure, whereas El-Gamal is quite slow when ciphertext modulus increases. On the other hand, the presentation between evaluated FHE schemes and CKKS is inconsistent, especially for multiplication timings. For both BFV and BGV, SEAL and PALISADE demonstrate a slower multiplication compared to CKKS. In contrast, HElib makes that of CKKS be a time-consuming process in comparison to the other two. Besides, between BFV and BGV, performance analysis has shown that the former is performing better than the latter in terms of execution time for key generation in both SEAL and PALISADE, mostly when the ciphertext dimension is climbed up.

To have a comprehensive survey on HE, in the next chapter, we will discuss the security of homomorphic cryptosystems in general and the security of these six schemes in particular under several security notions, such as IND-CPA and IND-CCA.

Chapter 5

On the Security of Homomorphic encryption

5.1 HE under security notions

All four general-purpose libraries presented in our work were based on RLWE-based systems. The most interesting advantage of LWE or RLWE is that it is considered as one of the hardest problems to solve in practical time for even post-quantum algorithms [LPR10]. However, this does not mean that RLWE-based HEs are totally secure. In fact, to prove security of encryption algorithms, two security models commonly referred are IND-CPA and IND-CCA, standing for *Indistinguishability under chosen plaintext attack* and *Indistinguishability under chosen ciphertext attack* respectively. For IND-CCA, there are IND-CCA1 and IND-CCA2. The former is Indistinguishability under non-adaptive chosen ciphertext attack, while the latter is the adaptive one.

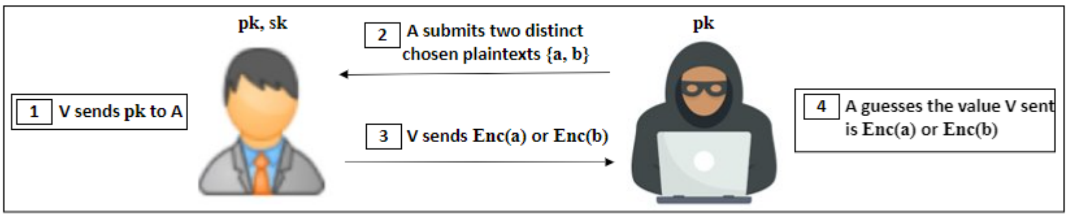

FIGURE 5.1: IND - CPA security notion.

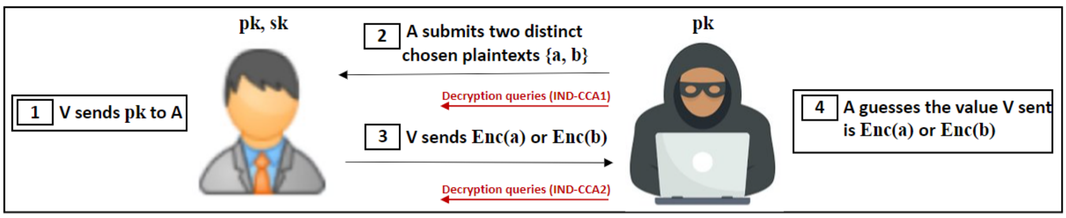

FIGURE 5.2: IND - CCA security notion.

IND-CPA is modeled by a game between an adversary (A) and a verifier (V) as illustrated in Figure 5.1. In general, after generating pk, sk, and other security parameters of an encryption system, V sends pk to A. From this point, A is free to perform any computations using pk. A then chooses two different plaintexts a, b and send them to V. V computes encryption of a or b uniformly at random and sends

A the result, called *challenge*. Finally, A needs to conclude the received value is the encryption of a or b. The cryptosystem is said to be secure in terms of IND-CPA if no adversary can output the correct value with probability significantly better than $\frac{1}{2}$. Likewise, the definition of IND-CCA is similar to IND-CPA, but here in both IND-CCA1 and IND-CCA2, the attacker can ask for the decryption of any ciphertexts, except the challenge that the verifier sent. In particular, IND-CCA1 and IND-CCA2 all allow the attacker to make queries to the decryption oracle to decrypt any arbitrary ciphertexts before the step 3 in Figure 5.1, when the verifier sends the challenge to the adversary. However, after the step 3, the adversary may not make further calls to the decryption device in IND-CCA1, while it is allowed in IND-CCA2 as illustrated on Figure 5.2. The security under IND-CCA2 implies the security under IND-CCA1, and the security under IND-CCA1 also implies the security under IND-CPA. In other words, an encryption scheme which is IND-CCA2 secure is both IND-CCA1 and IND-CPA secure.

5.1.1 RLWE - based FHE and SWHE schemes

In 1999, Bellare et al. [Bel+98] proved that all homomorphic encryption schemes are not secure against IND-CCA2 attacks. Subsequently, although IND-CCA2 is the strongest of the three security definitions, it is universally acknowledged that IND-CCA1 is the strongest security notion for HE. Apart from these three, Chenal and Tang [CT14] mentioned one variation of these security notions, called *key recovery* attacks. Following the authors, the key recovery attack is stronger than a typical IND-CCA1 and allows an adversary to recover the private keys through a number of decryption oracle queries.

Scheme	IND-CPA	IND-CCA1	Key recover attack
BFV			Z. Peng [Pen19]
BGV	✓	Fauzi et al. [FHR22]	Chenal and Tang [CT14]
CKKS			Li et al. [LM21]

TABLE 5.1: Security of several FHE and SWHE schemes.

Table 5.1 lists several FHE and SWHE schemes presented in our work and corresponding attacks, together with their related papers. It is obvious that three schemes are secure against IND-CPA attacks [FHR22]; however, they all suffers from IND-CCA1 and key recover attacks.

5.1.2 PHE schemes

In contrast to FHE and SWHE, one of PHE schemes, namely RSA, is weak even under IND-CPA norm. The reason is that Schoolbook RSA is deterministic. Therefore, comparing to IND-CPA model in Figure 5.1, to guess the correct output at step 4, the adversary can compute $a^e \pmod N$ and $b^e \pmod N$ then check which one is matched to the verifier's challenge. Thus, RSA is not IND-CPA secure, which also implies that it is not IND-CCA secure either. Unlike RSA, Paillier encryption is IND-CPA secure under Decisional Composite Residuosity (DCR) Assumption [GCD20]. To be more precise, Armknecht et al. [AKP13] proved that Paillier scheme is secure against IND-CCA1 attacks if and only if DCR^{SCCR} is hard, where SCCR is Subgroup Computational Composite Residuosity problem [Pai99]. Similar to Paillier system, El-Gamal encryption scheme is also known as being IND-CPA secure under the decisional Diffie-Hellman assumption [TY98]. However, when discussing the security of

El-Gamal under IND-CCA1, Wu and Stinson [WS08] supposed that it is conjectured, but there has been no formal proof.

Integer factorization problem (IFP). Apart from attacks on security notions, IFP is also worth discussing in our context when we have the security of both RSA and Paillier cryptosystems depending on factoring problem. Given a composite number N, The IFP is defined as finding two integers p and q such that $pq = N$. Once p and q are discovered, it can be shown that RSA and Paillier encryption are insecure (see chapter 3). Two of widely used algorithms to factor an integer, as well as to be the basis of other factorization methods, are Pollard's *rho* and Pollard's $p - 1$, invented by John Pollard in 1974 - 1975 [Pol74]. Our implementation and experimental results of each method are presented in detail at [DNT22].

With complexity of time and space $\mathcal{O}(\sqrt{N})$ by the birthday paradox, Pollard's *rho* relies on several important mathematical concepts, one of them is cycle-finding algorithm.

Algorithm 1: Pollard's *rho* algorithm using Floyd's cycle detection.

Input: a composite number N, a bound B for the number of iterations
Output: a nontrivial factor of N or failure

```
x ← 2 ;                          // Set x = x_0 = 2 to be the initial value
y ← 2 ;                          // Set y = x_0 = 2
d ← 1 ;
i ← 0 ;
while d = 1 OR d = N do
    if i ≥ B then
        return failure;          // Maximum number of iterations reached
    end
    x ← f(x) ;                   // x = x_i
    y ← f(f(y)) ;                // y = x_2i
    d ← gcd(|x − y|, N) ;
    i ← i + 1 ;
end
if d = 1 OR d = N then
    return failure ;
end
else
    return d ;
end
```

To reduce the memory cost, Pollard applied the idea of Floyd's cycle detection algorithm (see algorithm 1): two pointers x and y are used; pointer x holds the values of x_i's and pointer y holds the values of x_{2i}'s. Each iteration updates the values of x and y by computing $f(x)$ and $f(f(y))$, then checks if $\gcd(x_i - x_{2i}, N) = \gcd(x - y, N)$ is a nontrivial factor of N. This reduces the memory cost to $\mathcal{O}(1)$. Assuming f is a random function, then the expected number of evaluations to the function f performed by Pollard's *rho* algorithm is $\mathcal{O}(\sqrt{p}) = \mathcal{O}(\sqrt[4]{N})$, where p is the smallest prime factor of N. We present experimental results of Pollard's rho algorithm on our classical laptop with the following three numbers:

$$\begin{aligned} n_1 &= 1125939825397831601 \\ &\quad \text{(a 60-bit RSA modulus)} \\ n_2 &= 925276410789441750962080530947 \\ &\quad \text{(a 100-bit RSA modulus)} \\ n_3 &= 115792089237316195423570985008687907853269984 \\ &\quad 665640564039457584007913129639937 \\ &\quad \text{(Fermat number } F_8 = 2^{2^8} + 1) \end{aligned}$$

For each run we use the same function $f(x) = x^2 + 1 \mod N$ and same maximum number of iterations $B = 10^8$. The results are in Table 5.2 [DNT22].

Input number	Average running time (s)	Standard deviation
n_1	0.0054	0.0016
n_2	5.9370	3.3053
n_3	67.0664	43.6244

Table 5.2: Average running time with different initial values.

Table 5.3 shows the running time and size of factor when we run the algorithm with $B = 10^{10}$ to find a medium-size factor (around 20 digits) of some worst-case numbers:

$$\begin{aligned} n_4 &= 237130450584081431781941097598542348001 \\ &= 15351399207396244631 \times 15446829789289363271 \\ &\quad \text{(a 128-bit RSA modulus)} \\ n_5 &= 304075290252258958535257891241265214597 \\ &= 18229633569899862109 \times 16680274405204585033 \\ &\quad \text{(a 128-bit RSA modulus)} \\ n_6 &= 34! - 1 \\ &= 295232799039604140847618609643519999999 \\ &= 10398560889846739639 \times 28391697867333973241 \end{aligned}$$

Input number	Running time (s)	Digits in factor
n_4	970	20
n_5	819	20
n_6	27.1042	20

Table 5.3: Running time to find medium-size factor.

It can be seen that Pollard's *rho* takes a lot of time to find a factor of medium size. The second method to factor N is Pollard's $p-1$ algorithm, which is based on Fermat's Little Theorem [Pol74].

Unlike the previous method, the possibility of finding a factor p of given size is not determined solely by its size, but rather by the smoothness of $p-1$ (see algorithm 2). There are two cases of failure: $d = 1$ or $d = N$. In the first case, $a^M - 1$ is co-prime

Algorithm 2: Pollard's $p-1$ algorithm

Input: a composite number N, a bound B

Output: a nontrivial factor of N or failure

```
1  Choose a positive integer base a randomly between 1 and N;
2  Compute d = gcd(a, N);
3  if d ≠ 1 then
4  |  return d;
5  end
6  for prime numbers p_i ≤ B do
7  |  q ← 1;
8  |  while q ≤ B do
9  |  |  a ← a^{p_i} mod N;
10 |  |  q ← q × p_i;
11 |  end
12 |  c ← a − 1;
13 |  d ← gcd(c, N);
14 |  if d ≠ 1 AND d ≠ N then
15 |  |  return d;
16 |  end
17 |  if d = N then
18 |  |  Go to line 1 and choose a new value for a;
19 |  end
20 end
21 if d = 1 then
22 |  return failure;
23 end
```

with N, which implies that the search bound B is too small, and thus one should rerun the algorithm with a larger B. In the second case, $d = N$ implies that N has a B-smooth prime factor p, but the randomized base a has order less than $p-1$ modulo p (hence omitted for gcd computation in the loop from line 8 to 11). In this case we choose another base a and restart the whole process. Here, to improve the algorithm's performance, we implement the *two-stage* variant of Pollard's $p-1$ (the detail is presented in [DNT22]). The second-stage is performed by choosing a second bound $B_2 > B$, normally $B_2 = 100B$. While Pollard's *rho* takes much time to find a factor of medium size, our implementation of the *two-stage* Pollard's $p-1$ is able to find larger factors of some record numbers listed in [Lor21]. The running time for each is displayed in Table 5.4, where :

$$n_7 = 2^{977} - 1$$
$$n_8 = 575\text{th Fibonacci number}$$
$$n_9 = 960^{119} - 1$$

Digits of factor	n	B_1	B_2	Time(s)
32	$2^{977}-1$	10^7	10^8	14.9582
34	575th Fibonacci number	10^7	10^8	14.3806
66	$960^{119}-1$	10^8	10^{10}	1076

TABLE 5.4: Our running time of record factors by Pollard's $p-1$.

We found the factor of each number as below:

$$
\begin{aligned}
p_1 &= 49858990580788843054012690078841 \\
&\quad (\text{32-digit factor of } n_7) \\
p_2 &= 7146831801094929757704917464134401 \\
&\quad (\text{34-digit factor of } n_8) \\
p_3 &= 672038771836751227845696565342450315062141551 5 \\
&\quad (\text{66-digit factor of } n_9)
\end{aligned}
$$

It is obvious that Pollard's $p-1$ is capable to find large factors of a composit N; however, in practice, when the system applies N from 2048 bits, it is not sufficient to find its large-size factors using Pollard's *rho* and Pollard's $p-1$ methods on a classical machine. Therefore, the invention of *Shor's algorithm* by Peter Shor, a quantum computer algorithm to solve IFP, marks an important milestone in the security of public-key cryptography systems.

5.2 Shor's quantum algorithm

Being developed in 1994, Shor's algorithm [Sho94] is one of the first quantum algorithms that demonstrated the advantage of quantum computers over classical ones. In general, the method allows to find prime decomposition of big integers in polynomial time, namely $\mathcal{O}((\log N)^3)$ time and $\mathcal{O}(\log N)$ space, given a sufficiently large quantum computer.

The basic idea of Shor's algorithm relies on period-finding problem. Given integers a and N, r is called the *period* of a modulo N if r is the smallest positive integer such that a^r-1 is a multiple of N, or a^r-1 is divisible by N. For example, given $a=7$ and $N=15$, its period is found as $r=4$, we have $7^4 = 1 \pmod{15}$. The name "period" comes from the fact that $a^{i+r} \pmod N = a^i a^r \pmod N = a^i \pmod N$ (because $a^r = 1 \pmod N$) for any integer i. Based on the period's property, we have $N|(a^r-1)$. If r is even, then $N|[(a^{r/2}-1)(a^{r/2}+1)]$. By computing $\gcd((a^{r/2}-1), N))$ and $\gcd((a^{r/2}+1), N)$, we can find the factors of N. In the whole process, a quantum algorithm is applied to compute the period r of a modulo N by using quantum Fourier transforms [Sho99], where a is a randomly chosen element.

So far, the largest numbers factored by Shor's algorithm are 51 and 85 by Geller and Zhou in 2013 using eight qubits [GZ13]. Before that, Vandersypen et al. [Van+01] in 2001 and Martín-López et al. [Mar+12] in 2012 also implemented Shor's algorithm to factor 15 and 21 respectively. The most recent paper was of Gidney et al. [GE21], published in 2021, which presented how to factor 2048-bit RSA integers using 20 million noisy qubits in 8 hours.

Authors	Year	Gates	Total qubits
Shor [Sho94]	1994	$\mathcal{O}(n^3 \log n)$	$\mathcal{O}(n)$
Beckman et al. [Bec+96]	1996	$\mathcal{O}(n^3)$	5n+1
Veldral et al. [VBE96]	1996	$\mathcal{O}(n^3)$	4n+3
Beauregard [Bea02]	2003	$\mathcal{O}(n^3 \log \frac{n}{\epsilon} \log \frac{1}{\epsilon})$	2n+3
Takahashi et al. [TK06]	2006	$\mathcal{O}(n^3 \log \frac{n}{\epsilon} \log \frac{1}{\epsilon})$	2n+2
Haner et al. [HRS16]	2016	$\mathcal{O}(n^3 \log n)$	2n+2
Gidney [Gid17]	2017	$\mathcal{O}(n^3 \log n)$	2n+1

TABLE 5.5: Different implementations of Shor's algorithms on IFP.

With the efficiency of quantum computers, the security provided by cryptosystems, which are based on IFP and discrete logarithmic problems (DLP), seems to be short-lived. Speaking of IFP, RSA and Paillier encryption are vulnerable against Shor's algorithm. In [Suo+20], Suo et al. indicate some implementations of Shor's algorithm over different quantum prototype computers, together with their number of qubits and quantum gate complexities, as shown on Table 5.5 (for IFP) and Table 5.6 (for DLP).

Authors	Year	Time complexity	Space complexity
Shor [Sho94]	1994	$\mathcal{O}(n^3)$	$\mathcal{O}(n)$
Proos et al. [PZ03]	2003	$\mathcal{O}(n^2)$	-
Ekera et al. [Eke19]	2019	–	$\mathcal{O}(n^2)$

TABLE 5.6: Different implementations of Shor's algorithms on DLP.

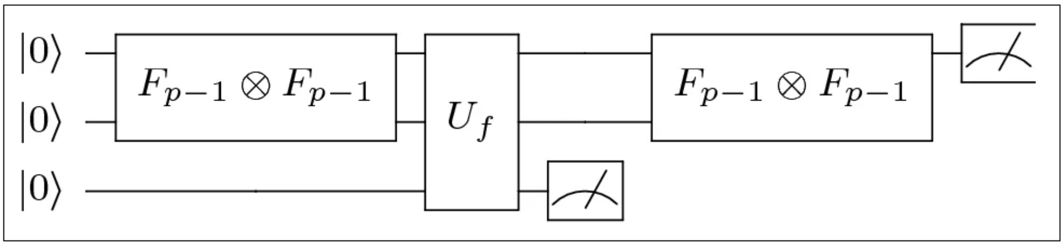

FIGURE 5.3: A circuit of DLP.

Similarly, El-Gamal with its hardness of computing discrete logarithms is also a victim of quantum algorithms [Sho94]. In 2010, Wang [Wan10] defined a circuit for quantum computers to solve DLP as shown on Figure 5.3, where F_{p-1} is the Fourier transform over Z_{p-1}, and U_f being a quantum circuit.

Some argue that although quantum encryption breaking is a potential possibility, it is not a peril as there are still solutions for it. One is to increase the bit lengths, so that attackers need a larger and larger quantum computer to be able to successfully break the system. The second is to develop new public key cryptosystems that cannot be solved by Shor's algorithm. This opens a new era of Post-quantum cryptography (PQC), or quantum-resistant cryptography.

Chapter 6

Conclusions and Future work

Nowadays, in the vibrant and active world, when data privacy plays a more significant role, *homomorphic encryption* (HE) is a new promising domain that allows external third parties to perform computations on the encrypted data without decrypting it in advance. However, one big challenge is to build a HE scheme that provides simultaneously both the required security and the efficiency in implementation.

In this work, we contributed an in-depth study of the different uses and implementations of HE schemes in most-used HE libraries, including SEAL, PALISADE, HElib, and HEAAN. Being different to previously related papers, six presented HE schemes cover all three homomorphic encryption categories: Partially HE, Somewhat HE, and Fully HE. First of all, we highlighted the principles and mathematical models of adopted schemes, followed by a brief description of linked libraries. Secondly, by comparing execution time of five main homomorphic operations (`KeyGen`, `Enc`, `Dec`, `Add`, `Mult`), we gave a computational overview of performance evaluation of different HE cryptosystems in different libraries. In fact, choosing a suitable encryption scheme and an appropriate HE library for it in a real application are not a simple problem. At least two foremost factors to consider are the data type and operations which users needed to perform in the context. For example, if users want to do only one type of computational operations, such as only addition or multiplication, then the a PHE is the best solution. In other cases, if they want to perform several types of computations at the same time on a ciphertext, then a SWHE or FHE is the better choices, and so on. Therefore, our experimental results, together with our hands-on implementation analysis, aim at making things easier for developers, especially non-experienced practitioners, to set input parameters for building their own HE-based projects. Additionally, we also demonstrated an overview of the security of aforementioned HE schemes under notable security notions such as IND-CPA, IND-CCA1 and IND-CCA2. Based on implementation of two classical attacks on Integer factorization problem, we discussed Shor's quantum algorithm for the same problem.

It is also clear that the efficiency of the PHEs become crucial in the overall performance. Due to the fact that PHEs are not implemented in mentioned libraries, we used our own implementations of Paillier, El-Gamal, and RSA as partially homomorphic cryptosystems in the emulation. For that reason, we plan to continue our work on optimizing PHE schemes in their implementation and performance. Besides, working with different HE libraries, we have seen that a part from doing encryption between a typical two parties, some libraries also support threshold encryption. A threshold cryptosystem allows n parties to communicate in which a minimal number of parties - a "threshold" number - need to cooperate in order to decrypt a ciphertext. This prevents the situation where an individual keyholder is able to decrypt all sensitive information on his own. This aspect will certainly be addressed by future work.

Bibliography

[Pol74] John M Pollard. "Theorems on factorization and primality testing". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 76. 3. Cambridge University Press. 1974, pp. 521–528.

[PH78] Stephen Pohlig and Martin Hellman. "An improved algorithm for computing logarithms over GF (p) and its cryptographic significance (corresp.)" In: *IEEE Transactions on information Theory* 24.1 (1978), pp. 106–110.

[RSA78] Ronald L Rivest, Adi Shamir, and Leonard Adleman. "A method for obtaining digital signatures and public-key cryptosystems". In: *Communications of the ACM* 21.2 (1978), pp. 120–126.

[ElG85] Taher ElGamal. "A public key cryptosystem and a signature scheme based on discrete logarithms". In: *IEEE transactions on information theory* 31.4 (1985), pp. 469–472.

[Mon94] Peter L Montgomery. "A survey of modern integer factorization algorithms". In: *CWI quarterly* 7.4 (1994), pp. 337–366.

[Sho94] Peter W Shor. "Algorithms for quantum computation: discrete logarithms and factoring". In: *Proceedings 35th annual symposium on foundations of computer science*. Ieee. 1994, pp. 124–134.

[Bec+96] David Beckman et al. "Efficient networks for quantum factoring". In: *Physical Review A* 54.2 (1996), p. 1034.

[VBE96] Vlatko Vedral, Adriano Barenco, and Artur Ekert. "Quantum networks for elementary arithmetic operations". In: *Physical Review A* 54.1 (1996), p. 147.

[Bel+98] Mihir Bellare et al. "Relations among notions of security for public-key encryption schemes". In: *Annual International Cryptology Conference*. Springer. 1998, pp. 26–45.

[TY98] Yiannis Tsiounis and Moti Yung. "On the security of ElGamal based encryption". In: *International Workshop on Public Key Cryptography*. Springer. 1998, pp. 117–134.

[Pai99] Pascal Paillier. "Public-key cryptosystems based on composite degree residuosity classes". In: *International conference on the theory and applications of cryptographic techniques*. Springer. 1999, pp. 223–238.

[Sho99] Peter W Shor. "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer". In: *SIAM review* 41.2 (1999), pp. 303–332.

[Van+01] Lieven MK Vandersypen et al. "Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance". In: *Nature* 414.6866 (2001), pp. 883–887.

[Bea02] Stephane Beauregard. "Circuit for Shor's algorithm using 2n+ 3 qubits". In: *arXiv preprint quant-ph/0205095* (2002).

[PZ03] John Proos and Christof Zalka. "Shor's discrete logarithm quantum algorithm for elliptic curves". In: *arXiv preprint quant-ph/0301141* (2003).

[BGN05] Dan Boneh, Eu-Jin Goh, and Kobbi Nissim. "Evaluating 2-DNF formulas on ciphertexts". In: *Theory of cryptography conference*. Springer. 2005, pp. 325–341.

[TK06] Yasuhiro Takahashi and Noboru Kunihiro. "A quantum circuit for Shor's factoring algorithm using 2n+ 2 qubits". In: *Quantum Information & Computation* 6.2 (2006), pp. 184–192.

[FG07] Caroline Fontaine and Fabien Galand. "A survey of homomorphic encryption for nonspecialists". In: *EURASIP Journal on Information Security* 2007 (2007), pp. 1–10.

[WS08] Jiang Wu and Douglas R Stinson. "On the security of the ElGamal encryption scheme and Damgard's variant". In: *Cryptology ePrint Archive* (2008).

[Gen09] Craig Gentry. *A fully homomorphic encryption scheme*. Stanford university, 2009.

[Reg09] Oded Regev. "On lattices, learning with errors, random linear codes, and cryptography". In: *Journal of the ACM (JACM)* 56.6 (2009), pp. 1–40.

[LPR10] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. "On ideal lattices and learning with errors over rings". In: *Annual international conference on the theory and applications of cryptographic techniques*. Springer. 2010, pp. 1–23.

[Wan10] Frédéric Wang. "The hidden subgroup problem". In: *arXiv preprint arXiv:1008.0010* (2010).

[Bra12] Zvika Brakerski. "Fully homomorphic encryption without modulus switching from classical GapSVP". In: *Annual Cryptology Conference*. Springer. 2012, pp. 868–886.

[FV12] Junfeng Fan and Frederik Vercauteren. "Somewhat practical fully homomorphic encryption". In: *Cryptology ePrint Archive* (2012).

[Mar+12] Enrique Martin-Lopez et al. "Experimental realization of Shor's quantum factoring algorithm using qubit recycling". In: *Nature photonics* 6.11 (2012), pp. 773–776.

[AKP13] Frederik Armknecht, Stefan Katzenbeisser, and Andreas Peter. "Group homomorphic encryption: characterizations, impossibility results, and applications". In: *Designs, codes and cryptography* 67.2 (2013), pp. 209–232.

[GZ13] Michael R Geller and Zhongyuan Zhou. "Factoring 51 and 85 with 8 qubits". In: *Scientific reports* 3.1 (2013), pp. 1–5.

[Hea+13] James Heather et al. "Solving the Discrete Logarithm Problem for Packing Candidate Preferences". In: *International Conference on Availability, Reliability, and Security*. Springer. 2013, pp. 209–221.

[BGV14] Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. "(Leveled) fully homomorphic encryption without bootstrapping". In: *ACM Transactions on Computation Theory (TOCT)* 6.3 (2014), pp. 1–36.

[BV14] Zvika Brakerski and Vinod Vaikuntanathan. "Efficient fully homomorphic encryption from (standard) LWE". In: *SIAM Journal on computing* 43.2 (2014), pp. 831–871.

[CT14] Massimo Chenal and Qiang Tang. "On key recovery attacks against existing somewhat homomorphic encryption schemes". In: *International Conference on Cryptology and Information Security in Latin America*. Springer. 2014, pp. 239–258.

[HS14] Shai Halevi and Victor Shoup. "Algorithms in helib". In: *Annual Cryptology Conference*. Springer. 2014, pp. 554–571.

[LN14] Tancrede Lepoint and Michael Naehrig. "A comparison of the homomorphic encryption schemes FV and YASHE". In: *International Conference on Cryptology in Africa*. Springer. 2014, pp. 318–335.

[Par+14] Payal V. Parmar et al. "Survey of various homomorphic encryption algorithms and schemes." In: *International Journal of Computer Applications 91* (2014).

[HRS16] Thomas Häner, Martin Roetteler, and Krysta M Svore. "Factoring using 2n+ 2 qubits with Toffoli based modular multiplication". In: *arXiv preprint arXiv:1611.07995* (2016).

[MBF16] Vincent Migliore, Guillaume Bonnoron, and Caroline Fontaine. "Determination and exploration of practical parameters for the latest Somewhat Homomorphic Encryption (SHE) Schemes". In: (2016).

[Che+17] Jung Hee Cheon et al. "Homomorphic encryption for arithmetic of approximate numbers". In: *International Conference on the Theory and Application of Cryptology and Information Security*. Springer. 2017, pp. 409–437.

[Gid17] Craig Gidney. "Factoring with n+ 2 clean qubits and n-1 dirty qubits". In: *arXiv preprint arXiv:1706.07884* (2017).

[Lai17] Kim Laine. *Simple encrypted arithmetic library 2.3.1*. https://www.microsoft.com/en-us/research/uploads/prod/2017/11/sealmanual-2-3-1.pdf. 2017.

[MSM17] Paulo Martins, Leonel Sousa, and Artur Mariano. "A survey on fully homomorphic encryption: An engineering perspective". In: *ACM Computing Surveys (CSUR)* 50.6 (2017), pp. 1–33.

[Aca+18] Abbas Acar et al. "A survey on homomorphic encryption schemes: Theory and implementation". In: *ACM Computing Surveys (Csur)* 51.4 (2018), pp. 1–35.

[HS18] Shai Halevi and Victor Shoup. "Faster homomorphic linear transformations in HElib". In: *Annual International Cryptology Conference*. Springer. 2018, pp. 93–120.

[DH19] Whitfield Diffie and Martin E Hellman. "New directions in cryptography". In: *Secure communications and asymmetric cryptosystems*. Routledge, 2019, pp. 143–180.

[Eke19] Martin Ekerå. "Revisiting Shor's quantum algorithm for computing general discrete logarithms". In: *arXiv preprint arXiv:1905.09084* (2019).

[Mas+19] Oliver Masters et al. "Towards a homomorphic machine learning big data pipeline for the financial services sector". In: *Cryptology ePrint Archive* (2019).

[Pen19] Zhiniang Peng. "Danger of using fully homomorphic encryption: A look at Microsoft SEAL". In: *arXiv preprint arXiv:1906.07127* (2019).

[RLF19] VF Rocha, Julio López, and V Falcão Da Rocha. *An Overview on Homomorphic Encryption Algorithms*. 2019.

[Yon19] Song Yongsoo. "Introduction to CKKS". In: Private AI Boot-camp, Microsoft Research, 2019.

[ALM20] Bechir Alaya, Lamri Laouamer, and Nihel Msilini. "Homomorphic encryption systems statement: Trends and challenges". In: *Computer Science Review* 36 (2020), p. 100235.

[De 20] Leo Ramón Nathan De Castro. "Practical homomorphic encryption implementations & applications". PhD thesis. Massachusetts Institute of Technology, 2020.

[GCD20] Ying Guo, Zhenfu Cao, and Xiaolei Dong. "A Generalization of Paillier's Public-Key System With Fast Decryption". In: *Cryptology ePrint Archive* (2020).

[HEl20] HElib v2.2.1. https://github.com/homenc/HElib. IBM, 2020.

[Suo+20] Jingwen Suo et al. "Quantum algorithms for typical hard problems: a perspective of cryptanalysis". In: *Quantum Information Processing* 19.6 (2020), pp. 1–26.

[YG20] Wei Yuan and Han Gao. "An Efficient BGV-type Encryption Scheme for IoT Systems". In: *Applied Sciences* 10.17 (2020), p. 5732.

[Che+21] Jung Hee Cheon et al. *Implementation of HEAAN*. https://github.com/snucrypto/HEAAN. 2021.

[GE21] Craig Gidney and Martin Ekerå. "How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits". In: *Quantum* 5 (2021), p. 433.

[HC21] Subir Halder and Mauro Conti. "Crypsh: A novel iot data protection scheme based on bgn cryptosystem". In: *IEEE Transactions on Cloud Computing* (2021).

[KPZ21] Andrey Kim, Yuriy Polyakov, and Vincent Zucca. "Revisiting homomorphic encryption schemes for finite fields". In: *International Conference on the Theory and Application of Cryptology and Information Security*. Springer. 2021, pp. 608–639.

[LM21] Baiyu Li and Daniele Micciancio. "On the security of homomorphic encryption on approximate numbers". In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2021, pp. 648–677.

[Lor21] Loria. *Record factors found by Pollard's $p-1$ method*. https://members.loria.fr/PZimmermann/records/Pminus1.html. 2021.

[pyb21] pybind11. https://github.com/pybind/pybind11. 2021.

[Zar+21] Christiana Zaraket et al. "Cloud based private data analytic using secure computation over encrypted data". In: *Journal of King Saud University-Computer and Information Sciences* (2021).

[FHR22] Prastudy Fauzi, Martha Norberg Hovd, and Håvard Raddum. "On the IND-CCA1 Security of FHE Schemes". In: *Cryptography* 6.1 (2022), p. 13.

[MT22] Saja J Mohammed and Dujan B Taha. "Performance Evaluation of RSA, ElGamal, and Paillier Partial Homomorphic Encryption Algorithms". In: *2022 International Conference on Computer Science and Software Engineering (CSASE)*. IEEE. 2022, pp. 89–94.

[Pol+22] Yuriy Polyakov et al. "Palisade lattice cryptography library user manual". In: *Cybersecurity Research Center, New Jersey Institute ofTechnology (NJIT), Tech. Rep* (2022).

[SWN22] Vasily Sidorov, Ethan Yi Fan Wei, and Wee Keong Ng. "Comprehensive Performance Analysis of Homomorphic Cryptosystems for Practical Data Processing". In: *arXiv preprint arXiv:2202.02960* (2022).

[Doa22] Thi Van Thao Doan. *Implementation of PHE schemes: El-Gamal, Paillier and RSA*. https://github.com/ThaoDoanVan/PHE. May 2022.

[DNT22] Thi Van Thao Doan, Thi Mai Phuong Nguyen, and Danh Nam Tran. *Simple Methods for Factorization*. https://github.com/ThaoDoanVan/Factorization. Project report. Sciences and Technologies Faculty, University of Limoges, Jan. 2022.

[PAL20] PALISADE v1.10.6. https://gitlab.com/palisade/palisade-release. PALISADE Project, Dec. 2020.

[SEA22] SEAL - Python. https://github.com/Huelse/SEAL-Python. Microsoft SEAL 4.X For Python, May. 2022.

[SEA20] SEAL (release 4.0). https://github.com/microsoft/SEAL. Microsoft Research, Redmond, WA, Apr. 2020.