Post Quantum Cryptography

Alternative Solutions to Hard Problems for Security

Kelechi Chukwunonyerem Emerole

A Thesis submitted for the Award of Doctor of Philosophy in Electrical, Electronic and Computer Engineering

July 2023

Post Quantum Cryptography

Alternative Solutions to Hard Problems for Security

Kelechi Chukwunonyerem Emerole

A Thesis submitted for the Award of Doctor of Philosophy in Electrical, Electronic and Computer Engineering

July 2023

Certificate of Authorship

I, Kelechi Chukwunonyerem Emerole, declare that this thesis titled, Alternative Solutions to Hard Problems in Post Quantum Cryptography and the work presented in it are my own under supervision. I confirm that:

- This work was done wholly or mainly while in candidature for a Doctor of Philosophy degree at Newcastle University, United Kingdom.
- Where any part of this thesis has previously been submitted to conferences or archived in a pre-print database, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed and the source given.
- All sources from software repositories have been acknowledged.

Signed: 

Date: ***17/05/2023***

Dedication

I dedicate this PhD thesis to my Lord and Saviour Jesus Christ, my wife Dr Mrs Laura Emerole, my daughter Joan Emerole, my son Michael Emerole and my mother Dr Mrs Chioma Emerole.

Acknowledgments

I want to acknowledge the guidance and assistance of my supervisor Prof Said Boussakta whose immense help and support glided me through this research journey. I also appreciate the leadership of my manager Dr Charles Morisset when I took up job as a Teacher in the university and the willingness of my teachers Dr Masoud and Dr Giacomo Bergami to proof read my Thesis even on short notice. I also want to acknowledge my sponsors Federal Polytechnic Nekede on behalf of the Tertiary Education Trust Fund, the rectors of the institution; Dr Mrs C. Njoku and Dr Michael Arimanwa for their approval to go for studies. In addition, I thank the Head of Department, Electrical and Electronic Engineering, Federal Polytechnic Nekede, Engr. Maurice Anyaehie for covering me during my absence. I also acknowledge staff of the School of Electrical Engineering and School of Computing, Newcastle University, United Kingdom for their support throughout my studies. I also appreciate the work of the examiners Prof Gui Tan and Prof Fatih Kurugollu. Finally, I appreciate the effort of my family in their guidance and purposeful discussions towards the success of the programme, my wife Dr Mrs Laura Emerole, my sister Chiagoziem Emerole, My mother Dr Mrs Chioma Emerole, my father Dr Chima Emerole and the ideas from my colleagues Dr. Moses Nsidibe-Obong, Dr. Cavus Mohammed that made the programme a resounding success.

Abstract

Post Quantum cryptography are defined as public key crypto algorithms whose public keys are generated from hard computational problems that are complex to solve in polynomial time by a quantum computer given worst case instances. The hard problems which have been proven to be quantum resistant include the shortest vector problem of lattices, the syndrome decoding problem of certain error correcting codes and the isomorphism of polynomial problem of multivariate quadratic polynomials. Solutions to these problems have been proposed which in turn have impact on the security and storage cost of such algorithms to protect information systems in the future. In this thesis, alternative solutions are proposed which are based on robust and complex vector space mappings. Firstly, Dimensionality mapping is proposed to reduce the basis into its linear independent vectors at low dimensionality by constructing a collapse function as an optimization problem. This optimization problem can be solved on the condition that a projection of the basis vectors from the High dimensional space to low dimensional manifold would have nearly orthogonal constitution. These eliminates the need for pre-processing using Gram-Schmidt Orthogonalization process. Implementing this approach on a channel basis, showed an improved BER performance over the Lenstra-Lenstra-Lovatsz algorithm for about 1db and 4db in the 4×4 and 6×6 uncoded system using 4QAM constellation. Secondly, the solution of the syndrome decoding problem is generalized to codes associated with the totally non-negative Grassmannian. The solution was reduced to an instance of finding a subset of the Plücker coordinates with the minimum Grassmann distance from the subspace containing the encrypted message symbols. Furthermore, bounds where derived which showed that the complexity scales up on the size of the Plücker coordinates. In addition, experimental results on decoding failure probability and complexity based on row operations where presented and compared to Low Density parity check codes in the Hamming metric. Finally, the kernel function of the New Mersenne number transform was applied to hide the

structure of the core map(central polynomial) of a multivariate polynomial based cryptosystem. This is in order to mitigate the interpolation of the rank of the quadratic form by an adversary. The implementation of this new isomorphism from the New Mersenne Number Transform showed an average of 69% reduction in secret key size. Further implementation of the isomorphism against key recovery attacks from the MinRank instance where carried out and it was shown that for lower field sizes the new isomorphism had an average success time of 13.8%.

Contents

List of Tables

List of Figures

List of Acronyms

SVP Shortest Vector Problem

SDP Syndrome Decoding Problem

HFE Hidden Field Equations

ISP Isomorphism of Polynomial

RSA Rivest Shamir Adleman

LLL Lenstra Lenstra Lovasz

NMNT New Mersenne Number Transform

LPDC Low Density Parity Check

LRPC Low Rank Parity Check

GS0 Gram Schmidt Orthogonalization

BER Bit Error Rate

LWE Learning with Error

SBL Sample Basis left

SBR Sample Basis Right

CPA Chosen Plaintext Attack

DRSD Decisional Rank Syndrome Decoding

MPDC Moderate Parity Density Check

RSL Rank Support Learning

CSD Computational Syndrome Decoding

ISD Information Set Decoding

GRS Gaborit Ruata Schrek

LLR Log Likelihood Ratio

SNR Signal to Noise ratio

BP Belief Propagation

LT Luby Transform

XL Extended Linearization

FE Fast Evaluation

UV Unbalanced Vinegar

GCE Gray Code Enumeration

MFE Medium Field Equations

ZHFE Zuang-Zi Hidden Field Equations

CMAC Macaulay Matrices Columns

RMAC Macaulay Matrices Rows

FPA Floating Point Arithmetic

AVX Advanced Vector Extensions

KL Kullback-Leibler

MIMO Massive Input Massive Output

QAM Quadrature Amplitude Modulation

BLISS Bimodal Lattice Signature Scheme

GVB Gilbert-Varshanov bound

AMD Advanced Micro Devices

BIAWGN Binary Input Additive Gaussian Noise

NP Non-deterministic Polynomial-time

NTT Number Theoretic Transform

Openssl Open source Secure Socket layer

Chapter 1

Introduction

1.1 Introduction

Classical computers provide solutions to hard problems on the input of a parameter that grows polynomially as function of the number of digits of the input. This computation is hastened with the introduction of random number generator imbued with succinct precision. The Quantum computer carries out computation within a 2^n dimensional vector space without the need of a random number generator. The basis of these vector space is made up of quantum states. However, its unique vector space mapping properties makes its solution to hard problems to pan out asymptotically with input size. These unique mapping properties can be generalized to the classical case as an alternative to the classical approaches proposed in literature. The kernel of the mapping is represented by a unitary matrix which is constructed using discrete Fourier transform. This is analogous to the kernel of the isomorphism used in Multivariate cryptography [1]. These mapping vis a vis transformation centred solutions to problems inspired the ideas behind the algorithms proposed in this PhD thesis. In addition, hard problems are randomized to its instance, that is reduced to an instance of the bigger problem. A solution to an instance will ultimately lead to a solution to the bigger problem albeit with negligible error. In other words, the solution is an approximation and not an exact solution. For example, to break the RSA cryptosystem which is based on the hardness of factorizing a prime number, firstly, the problem is reduced to the problem of finding the order of an element in the multiplicative group (modular of the prime number) [2]. Subsequently, to find a factor of the prime number with certain probability, the greatest common divisor of an exponential of the element as a function

of the order and the prime number is computed.

Hard computational problems are the mathematical requirements necessary to construct the key generation scheme of cryptosystem. This is because they are the central ingredient necessary to generate the two keys; public key and private key used in Public key cryptography. It has been proved from research that classical mathematical problems such as the Shortest Vector Problem for lattices, Syndrome Decoding problem for codes and Isomorphism of polynomial problem from Hidden Field equations are strong against Shor's Quantum based algorithm [2]. This algorithm hitherto has been successful in solving the Integer factorization problem of the RSA and the discrete logarithm problem of Elliptic curve cryptography. These algorithms have finally passed standardization and is expected to be deployed into the Openssl framework for data encryption, decryption and authentication [3]. However, research on the cryptanalysis of these Post quantum algorithms are ongoing and this PhD thesis is building on that to expand the techniques used in analyzing these algorithms. One of the proposal is that the mathematical problem to be solved can be re-constructed into an optimization problem which can be solved by statistics based theoretic methods, graph based methods, machine learning based methods and signal processing methods. The performance of this approach would be compared with methods that have been employed in literature to solve the problem. This application moves away from the application-centred nature of some of these methods but rather extract its mathematical construction and re-tune them to serve the purpose of increasing the discourse in the subject matter till a robust mechanism is in place just in time for the deployment of a super quantum computer. The research is heavily analytical and this time consuming process has made hardware implementation practically impossible.

The first contribution of the PhD research was to propose a more efficient way of decomposing the basis in order to solve the shortest vector problem for lattices. The novel method employed was the use of dimensionality mapping [4]. Shortest Vector problem is an intractable mathematical basis for the design of Learning with error-based Lattice cryptography which is a cryptographic primitive that has the capability to secure data against quantum threats [5] [6]. The method was compared to Lenstra-Lenstra-Lovasz method [7] and applied to reduce a lattice basis with rank $\leq n$ where n is the dimension of the lattice. The result showed good bit error rate performance as result of the quality of the reduced basis. The major benefit of the proposed method is that it provides an efficient method of

cryptanalyzing lattice based cryptosystems with reduced parameters. The second contribution of the thesis was to generalize the syndrome decoding problem [8] [9] [10] to the Grassmann metric [11] [12] [13] [14] by applying the generalized information set decoding method [15] [16] [9] to find the Plücker coordinates containing the lowest coset of a totally non negative Grassmannian. Furthermore, the shortest vector problem was generalized to the syndrome decoding problem using the probability distribution as the metric for determining the coset with the lowest weight. Copula functions [17] [18], [19] were used to analyze these distributions. The major benefit of this is to advance codes that would be an alternative to Hamming weight based codes. The third contribution of the thesis was to apply the New Mersenne number transform [20] [21] as an isomorphic map in the Isomorphism of Polynomial problem. This was compared to the Affine transform used in literature [22], [23] [24]. The multivariate based cryptosystem are those constructed with Hidden field equations. The benefit of this is the reduction in computational complexity.

Lattices are linear, efficient in implementation and are based on the difficulty of finding the shortest vector in a lattice when given a reference Gaussian distribution. Many lattice based construction require what is called an SVP oracle [25] which is efficient and fast to sample lattice points given a basis and a vector. In this thesis, the SVP oracle of importance is the Gaussian sampler. They are used in generating random numbers as private keys for encryption and for security proof because of its ease of approximation. The lattice vector is distributed with a centre c and a standard deviation that is close to the centre. Given the basis, it is possible to produce the vector without disclosing information about the basis. Also [26] convolution of the probability density function for q-ary lattices has been employed to prove statistical indistinguishability. The Lenstra-Lenstra-Lovastz reduction method for lattices is basically a pre-code method to process the lattice which has a matrix structure that is oblivious to the polynomial time adversary. The lattice basis is decomposed into its Gram-Schmidt equivalent which is highly sparse and consumes so much memory. However, the theory of Dimensionality mapping is proposed which reduces the lattice into its linearly independent equivalent after some level of approximation. The mathematical properties of the noise generated from the Gaussian sampler to scramble the basis is interesting as well as the correctness of these sampler which can be proved using a convolution theorem. The complexity of Gaussian sampling can also be attributed to the standard deviation of the distribution where $\sigma \geq \omega(\sqrt{\log n}).\max 1 \leq i \leq n||B||$ where $||B||$ is

the norm of the basis [27]. To improve algorithmic complexity, the Gram-Schmidt basis should pre-computed and stored before the sampling process which will lead to $O(n^2)$ operations[28]. In the course of the research, it was discovered that there was a connection between the theory of lattices and the theory of error correcting codes. The lattice points can be modelled just as codewords. Proving the security of a lattice based construction is more efficient in terms of public key size, signature size and computation in the random oracle model than in the standard model[29]. Notably master public key sizes on input of a security parameter λ is exponentially equivalent to $O(\lambda)$. Furthermore, this can be transformed through a function that employs the tangential plane from the reference plane of the vector. These planes are actually orthogonal subspaces. This has been studied extensively by Klein who proposed a random function[30] which is an improvement of Babai's approach [31]. It is usual that a system samples from a distribution statistically close to the Gaussian distribution to within 2100 by employing floating point operations that have precision of at least 100 bits[32]. Using standard (53 bit) double precision floating point numbers is efficient as compared to the use of multi-precision arithmetic numbers, but not up to 80bit or 100bit security levels. Floating point arithmetic require pre-computed tables which increase computation time.

Code based cryptography which relies on hardness of decoding syndromes as its security metric has proven to be resilient to quantum attacks. The first real effort to formalize it was the technique by McEliece[33] which employed binary Goppa codes whose security metric relied on the hardness of decoding a linear code and the difficulty in differentiating Goppa codes from other random codes. Despite the advantages, the storage requirements in terms of the public key size are still an open problem. Research into applying some other family codes like the Quasi-cyclic codes, Low density parity check codes and a concatenation of both to solve this problem have been pursued vigorously[34]. The weight of the rows and advent of structural cryptanalysis have made them impractical for use. Also Quasi-cyclic Moderate density parity check code has also been proposed[35]. Using Low Rank Parity check codes has been proposed to reduce the decoding error probability. This employs the rank metric instead of the Hamming metric associated with Low Parity Density Codes. The Hamming metric specifies the number of vector coordinates that distinguishes one vector from another. The approach of using Rank Parity check codes was to mitigate against structural attacks which stems from the algebraic nature of syndromes. For a small key size and appropriate choice of parameters, an

equivalent of 280 bits of security can be achieved. Gabidulin introduced the rank metric and the Gabidulin codes over finite field with q^m elements, and constructed the first rank-based cryptosystem (GPT)[36] with much smaller key size compared to McEliece on Goppa codes. The syndrome decoding problem is defined as thus for a parity check matrix H , given an error vector e to find a syndrome of low hamming weight such that $He^T = s^T$. Efficient decoding of random linear codes in polynomial time is very important because it ensures practical implementation of code based cryptography. This is because knowledge of the cryptanalytic process through information set decoding methods can enhance the choice of parameters in order to enhance efficiency. The complexity of Chabaud and Stern[8] information set decoding method was reported to be $O(q(mr)(r1))$ and while that of Ourivski and Johansson [37] has an exponential term in $O(q(k+1)(r1))$ where n is the length of the code, k is the dimension and r is the rank.

Multivariate polynomials are made up of non-linear polynomials with several variables. These polynomials can be used to construct homogeneous equations. In order to solve problems from these polynomials, techniques like the Gröbner basis and its variant [38] are used to reduce the equations to linear equations that are solvable using sparse solutions. The hardness of the multivariate cryptosystem stems from the intractability of solving random non-linear homogeneous equations over finite fields and the isomorphism of polynomial problem [39] which is exponential on the structure of the map. These maps can either be bijective which depicts randomness or injective. The inversion of the injective map is more efficient than the inversion of the bijective map [40]. An example of the bijective map is the affine transform. These maps hide the structure of the core map from an adversary from making an inference and breaking the system. Patarin [22] developed an encryption scheme using Hidden Field equation after breaking the Matsumoto-Imai cryptosystem [41]. This involves a univariate polynomial map over the degree n extension field which generates a public key where the map is masked by two invertible transformations over a finite field. Patarin proposed a signature scheme called QUARTZ which was reported slow in performance because of the high degree over an extension field and was also bedevilled by inefficient invertible transformations [42]. Tao suggested that the degree should be $q^2 + 1$ where q is the cardinality of the finite field. The finite is ususally small for example $GF(2)$ [40]. This is in expense to large public key size from large design parameters.

1.2 Motivation and Problem Statement

One of the issues that mitigate the continued deployment of public key cryptography to secure information and communication infrastructure is the limitation of the mathematical and engineering tools used in breaking and analyzing the key security mechanism. The key security mechanism is based on Computational hard problems which increase in complexity for certain circumstances that depend on the parameters of the problem. Breaking classical public key cryptography algorithms using quantum mathematical tools has expanded the science and technology of cryptography to encompass principles from algebraic geometry, combinatorics, signal processing, optimization, discrete mathematics, polynomials, coding theory and so on. In some cases these limitations can have implication on the security of the cryptosystem and also key storage. Therefore the problems and limitations, this thesis seeks to investigate and proffer solutions are as follows:

- Complexity in decomposing the basis of the lattice which leads to complexity in computing the subset sum in the Gaussian sampling based solution of the shortest vector problem in lattice cryptography. This complexity can also be attributed to the limitations of pre-processing the orthornormal vectors using the Gram-Schmidt Orthogonalization process. The proposed solution which is explained in Chapter three is to use the theory of Dimensionality mapping in a manifold where an arbitrary projection is constructed which uses an affine transformation to map these points into a linearly independent subspace of lower dimensionality. This method eliminates the need for orthornormal vectors with its attendant computation overhead. Furthermore, to test the performance of this method in generating a high quality basis that can approximate this problem, the method was employed to reduce a Gaussian based channel basis and the BER result was compared with Lenstra–Lenstra–Lovász method. [7]. The sampling process is guided by the precision of the arithmetic used. Floating point arithmetic is recommended and bounds as a consequence of the security parameter have been derived in literature. However, in this thesis, there is further improvement to the bounds derived.

- Complexity in isomorphism from the isomorphism of polynomial problem for multivariate cryptography which leads to large private keys and longer time in generating signatures for authentication. This can be due to the dense nature

of the representative matrix of the inherent isomorphism. Also, limitation in isomorphism can lead to less robustness of the trapdoor function to key recovery attacks. The proposed solution is to use New Mersenne Number transform function as a masking function. The function has a kernel made up of primitive root of unity and imbued with an orthogonal property. Also, the transformed polynomial is invertible with the necessary condition that the inverse of the polynomial degree belongs to the ideals in the ring. The result was applied to the Gui cryptosystem [43] and the Sidon cryptosystem [44] and compared to the Affine transform.

- Limitations of information set decoding methods as a solution to the syndrome decoding problem for code based cryptography in the Grassmann metric. This is because in constructing the generator matrix of the code, the isomorphism of the system should be taken into account. The proposed solution was to derive a method for detecting the Plücker coordinate of totally non-negative Grassmanian. The Plücker coordinate are so chosen because the maximal minor is non-zero which makes it easier to construct the coordinates in a set. The failure probability of a solution to the problem was computed and compared to Low Density parity check codes. Furthermore, the sum over all transitional probabilities is not adequate to estimate the dependence of the positroid cells of the Non negative Grassmannian code. Therefore, copula functions were proposed in thesis to model the dependence of the Schubert cells of the Grassmannian in terms of subspaces.

1.3 Scope

The thesis studied the Computational hard problems and mathematical framework that are intractable for a quantum computer to solve. These mathematical problems have been employed to design the key generation method of certain crypto algorithms. The thesis also studied the mathematical solutions that have been proposed to solve these problems. The major aim of this thesis is to propose alternative solutions that will create a plethora of mathematical tools that would assist cryptanalysts to analyze the security of Post quantum public key cryptography algorithms in the future. These public key algorithms can withstand the threat of quantum computing capabilities. There are majorly two hard problems for encryption and two hard problems for authentication. The thesis studied three problems, shortest vector problem and syndrome decoding problem used in encryption and

multivariate quadratic problem used for authentication. The three solutions proposed in literature for the problems which are the LLL reduction, Information set decoding and Minrank solution which were also studied. The performance of the derived functions are tested with different programming languages; C/C++, Python, Sagemath and MATLAB. Solving the multivariate quadratic equation using Gröbner basis was beyond the scope of the Thesis. This is because this thesis targets solutions that borders on the key generation and recovery process and not algebraic attacks by means of reducing the core map to make it solvable by linearization. In addition, computing Gröbner basis involves studying the structure of Macaulay matrices which has tendency to stretch the thesis beyond its main focus considering the available time resource.

1.4 Publication

Substantial part of this work has been presented and published at conference venues and also archived in Arxiv pre-print platform.

- Kelechi Chukwunonyerem Emerole, Said Boussakta, Post Quantum Cryptography for IOT, Annual Research Conference,Newcastle University, Newcastle UK, 2019.

- Kelechi Chukwunonyerem Emerole, Said Boussakta, Optimizing Gaussian measure of lattices using Dimensionality Reduction, IEEE International Conference on Communications,Dublin Ireland, 2020. The beginning part of Chapter 3 was presented at this conference.

- Emerole, K. C., Said Boussakta (2021). Generalizing Syndrome Decoding problem to the totally Non-negative Grassmannian. arXiv preprint arXiv:2106. 15526. The beginning part of chapter 4 was archived at this pre-print portal.

- Kelechi Emerole, Optimizing information set decoding for the totally non negative Grassmanian, Postgraduate Research Conference, Newcastle University, Newcastle, UK, 2021.

- Kelechi Chukwunonyerem Emerole, Said Boussakta, Isomorphism in Multivariate Cryptography using the New Mersenne Number Transform, IEEE International Conference on Communications, Seoul South Korea, 2022. The beginning part of Chapter 5 was presented at this conference.

1.5 Chapter summary

The introduction of a random number generator with the appropriate precision enables the solution of NP problems. The vector spacing mapping capabilities of a quantum computer provides the tooling needed to solve hard problems without the need for random number generator. Hard problems can be solved by reducing the problem to an instance and using appropriate mapping of parameters to solve the instance of the bigger problem. Classical mathematical problems like the shortest vector problem in lattices of abelian groups, isomorphism of Polynomial problem of multivariate polynomials and syndrome decoding problem of error correcting codes are useful in designing the key generation algorithm of public key crypto algorithms that are presumed to be strong against attack by a quantum computer. These crypto algorithms can be attacked by reducing the solution to these problem as a solution to an optimization problem which is solvable using different techniques. The state-of-the-art techniques studied in this thesis to solve these problems include the Lenstra Lenstra Lovasz method for reducing lattice basis, Information set decoding methods for finding the coset weight of a codeword and the Affine isomorphism used in masking the core map of hidden field equations. The thesis seeks to propose alternative methods to solve these hard problems. Alternative approaches proposed in the thesis was briefly summarized in this chapter. For example, to decompose the basis of a lattice, dimensionality mapping was proposed, to generalize the solution to the syndrome decoding problem to the Grassmann metric, Plucker coordinate of the Grassmann based solution was proposed and finally to solve the isomorphism of polynomial problem, a new Isomorphism from the New Mersenne Number Transform was proposed. Due to the fact that this thesis focuses on key generation based attacks, discourse on algebraic attacks using Grobner basis was not carried out in this Thesis. This chapter ended with a record of the conference venues were this work was presented.

Chapter 2

Post Quantum Cryptography

In this chapter, the quantum resistant encryption construction from lattices, codes and hidden field equations are expounded. In addition, this chapter discusses concepts in lattice cryptography such as lattice basis, discrete Gaussian, basis reduction, basis delegation and a new concept of using triangularization to construct basis. The security analysis of such schemes as a consequence of games between an adversary and a challenger was also discussed in this chapter as well as information set decoding methods from linearized polynomials that form a set of quadratic equations. Finally, different constructions of hidden field equations for the core map of multivariate schemes was discussed as well.

2.1 Lattices

The advent of fast computation arising from the development of a quantum computer has made it imperative to develop algorithms that would mitigate sophisticated attacks to information systems due to the limitation of classical crypto algorithms based on intractable mathematical problems. One of such algorithms is constructed from lattices and it is based on the problem of finding the vector with the shortest Euclidean distance without knowing the structure of the basis. Lattices are linear and efficient in implementation. There are Ideal lattices which are parameters in a ring $\mathbb{Z}[x]/(x)\rangle$ for some irreducible polynomial f. They are useful in decreasing parameters for generating the lattice basis. They are also efficient in matrix arithmetic such as multiplication [45]. Despite these apparent advantages, it is less secured as compared to standard lattice based construction[46]. Security analysis of a lattice based construction is more efficient in terms of public key size, signature size and computation in the random oracle model than in the standard

model. In fact, on the input of a security parameter λ, master public key sizes is exponentially equivalent to (λ) [47]. Triangularization is a fast processing algorithm applied to matrix vector arithmetic which can be likened to subset sum in lattice cryptography. By reducing the structure of the lattice basis into its Toeplitz variant, it can be shown that operations on such a matrix leads to a quasi linear complexity. This is based on the assumption of [48] [49].

2.1.1 Review of works on lattice based encryption

A lattice based encryption was constructed using cover free families and the analysis of its security was done under the learning with error security assumption. The adversary in this system compromises a legitimate entity by possessing the decryption key [50]. The attack is carried out by querying the challenger to extract the keys for a bounded number of times. It is assumed that the system is secured even when the number of queries is increased. The system also employed Micciancio-Peikert's gadget matrix [51] in order to work with smaller parameters. The gadget matrix employed is sparse which makes storage inefficient and its inversion process has logarithmic complexity $O(n)$. Furthermore, a scheme was constructed using the subset difference method. This method is based on the technique of using binary trees and the security is analysed in the standard model. It is stated that the advantage of the subset difference method is that for a certain number of revoked leaf nodes r out of total leaf nodes $\mathbb{N}$ in a binary tree, the size of covering set is at most $2r - 1$ in the worst case scenario as compared to the covering set method[52]. This has a logarithmic complexity that increases the set from $(log\mathbb{N})$ to $(log^2\mathbb{N})$. However, when analyzing the security using games, a condition was stated that if a certain function $h_{id} = 0$, the simulator leaves the game and terminates the programme. It also leaves the game during its response to creating a ciphertext [53]. A hierarchical based broadcast encryption method was proposed using the Basis delegation approach. To reduce the complexity of the algorithm, node identities are deleted during encryption. This approach reduces the ciphertext to $(km+t)logq$ for positive integers k, m, t, q where $q \geq 2$ and $m \geq 2n\log q$. Using the indistinguishability from random oracle, the security of the scheme was analyzed against adaptive chosen-plaintext attacks and chosen identity attacks in the random oracle model [54]. A signcryption method was developed using the Basis delegation approach to maintain the dimension of the lattice and the scheme also employed Micciancio and Pierkert's trapdoor function. It was also proved that the scheme is unforgeable

against adaptive chosen message attacks under the small integer solution assumption. To reduce the complexity of the algorithm, node identities were deleted during encryption. This approach resulted in a ciphertext of size $k+l+(m+n[logq])$ [55]. A lattice based identity encryption with security proof against chosen identity and chosen message attack was constructed under the hardness of learning with errors assumption in the standard model. $l+1$ vectors were chosen to encode identities which served as the public key of the system. The basis delegation technique was not used. This is in order to extract keys from newly generated lattices which improves the efficiency of the system[56]. A lattice based identity encryption under the security assumption of learning with errors was proposed with a vector of length m. A function rot was used to construct matrices in order to reduce public key size from (mn) to (n) and to reduce the encryption complexity by employing a circular matrix with faster multiplication as the basis for the lattice[45]. Least parameter estimation was employed to generate a decodable lattice with least complexity as compared to the gadget basis whose sparse nature imposes serious storage cost.

2.1.2 Notation

Z is denoted as the set of integers and Z_q as set of integers modulo q in $(\frac{-q}{2}, \frac{q}{2})$ for any prime $q \geq 2$. $Z_q^{n\times m}$ is denoted as a set of $n \times m$ matrices with entries in Z_q. $poly(n)$ is denoted as the polynomial function of a security parameter n. For a large security parameter λ, a function $ngl : R \to R$ is negligible if $ngl(\lambda) \leq \frac{1}{P}(\lambda)$ for any polynomial $P(\lambda)$. B is denoted as a matrix and b as its column vector. Finally denote $(y_1, \ldots, y_m) \longleftarrow B(x_1, \ldots, x_m)$ as a function that takes the set of variables $(x_1, \ldots, x_m)$ as inputs and outputs the set of variables $(y_1, \ldots, y_m)$.

2.1.3 Lattice Basis

An n-dimensional full rank lattice $\Lambda \in Z^m$ is an additive subgroup of Z^m. A lattice is constructed by a basis $B \in Z_q^n$ with linearly independent vectors $b_1, \ldots b_n$ that generate the rows of B. For $B \in Z_q^n, u \in Z_q^m$ and where q is a prime and $m \geq 2n$, it can be written as [6]

$$\Lambda_q(B) = \{e \in Z^m, \exists s \in Z_q^n, B^T s = e(\mod q)\} \tag{2.1}$$

$$\Lambda_q^{\perp}(B) = \{e \in Z^m, Be = 0(\mod q)\} \tag{2.2}$$

$$\Lambda_q^{u}(B) = \{e \in Z^m, Be = u(\mod q)\} \tag{2.3}$$

In equation (2.1), the lattice points in an n-dimensional space are generated by the transposed rows of the basis, In equation (2.2) the vectors that generate the basis are orthogonal modulo q to its rows while in Equation (2.3), the bijective map is defined as $(e+\Lambda_{\perp}) \mapsto Be \mod q$. This means that decoding a codeword $Be \mod q$ is computed by reducing $emodulo\Lambda_{\perp}(B)$. In other words, for every codeword $u \in Z_q^n$ there is an error $e \in \{0,1\}^m$ such that the subset sum $Be = u \mod q$ where $m \geq 2n$. This is expounded further in the lemma 1[57]. The structure of a lattice is shown in Figure 2.1. The orthogonal coordinate vectors forms a basis of the lattice. It can also be seen that the orthogonal basis vectors have a shorter length than the basis vectors that are not orthogonal.

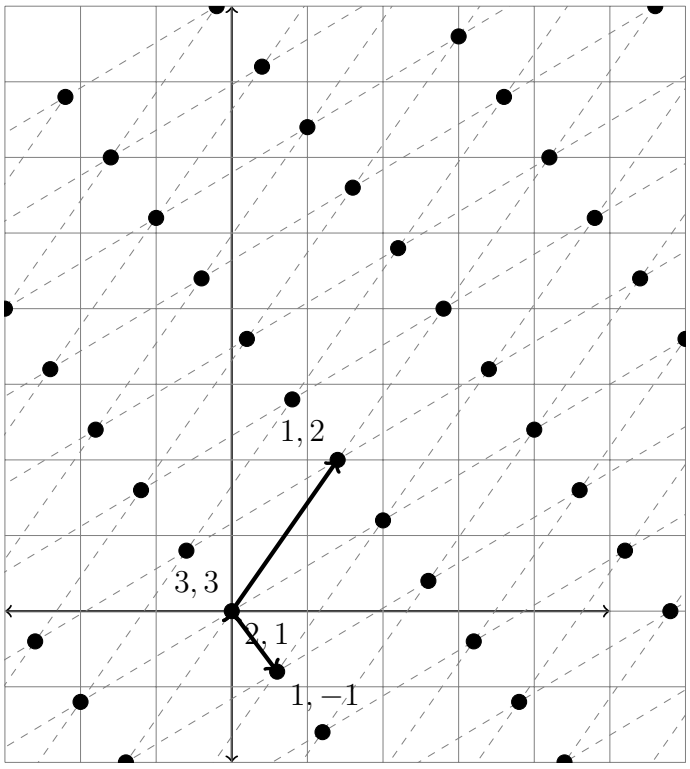

Figure 2.1: Dimension 2 Lattice with basis vectors at coordinate $(1, 2)$ and $(1, -1)$ orthogonal while basis vectors at coordinate $(2, 1)$ and $(3, 3)$ not orthogonal but linearly independent

Lemma 1. *The distribution $u = Be \mod q$ is within statistical distance 2ϵ of uniform distribution over Z_q^m, if the subset sum of column of $B \in Z_q^{n \times m}$ generate Z_q^n assuming $\epsilon \in (0, \frac{1}{2})$ and $\eta_{\epsilon}(\Lambda^{\perp}(B)) \geq s$ with the error sampled close to the distribution $D_{Z^m,s}$.*

Let $t \in Z^m$ be a solution to $Bt = u \mod q$ where $u \in Z_q^n$ is arbitrary, then the conditional distribution $e \sim D_{Z^m,s}$ given $Be \mod q$ is $t + D_{\Lambda^\perp,s,-t}$. The set of all syndromes $\{Be \mod q : e \in Z_m\} = Z_q^n$. By lemma 1, $e \sim D_{Z^m,s}$. It then holds that the distribution of $e \mod \Lambda^\perp$ is within statistical distance 2ϵ of distribution $(Z^m \setminus \Lambda^\perp)$. This distribution is isomorphic to Z_q^n via mapping $(e + \Lambda^\perp \mapsto Be \mod q)$. Writing $e = t + v$, and making a random choice $u \in Z_q^n$ with the e sampled from a distribution close to $D_{Z^m,s}$ and given the subset sum $Be = u \mod q$, the input to the distribution D is $t + \Lambda^\perp$ satisfies

$$D(e) = \frac{\rho_s(e)}{\rho_s(t + \Lambda^\perp)} = \frac{\rho_{s,-t}(e - t)}{\rho_{s,-t}(\Lambda^\perp)} = D_{\Lambda^\perp,s,-t}(e - t) \tag{2.4}$$

The Gram-Schmidt orthogonalization is also defined as $\tilde{B} = \{\tilde{b}_1, \ldots, \tilde{b}_n\} \subset Z^m$ where $r_i = \|\tilde{b}_i\|$ is the Euclidean norm. The bound on the Euclidean norm is summarized in the lemma 2

Lemma 2. *The Euclidean norm of a pair $|e^T x| \in (O, \frac{q}{2})$ satisfies $|e^T x| \leq \|e\|.(\alpha q.\omega(\sqrt{} + \frac{\sqrt{m}}{2})$ with high probability on input of m where $e \in Z^m$ and x is sampled from the distribution $\Psi_\alpha \in Z_q$.*

The error vector x is chosen randomly from a set of real numbers according to a random distribution with 0 mean and standard deviation $\frac{\alpha}{\sqrt{2\pi}}$ with random $\mod q$. The distribution of $Be = u(\mod q)$ is statistically close to uniform over Z_q^n where e is the sampled from the discrete Gaussian distribution. This holds for $3q^{-n}$ fractions of the basis $B \in Z_q^n$ and for the Gaussian parameter $\sigma \geq \omega(\sqrt{\log m})$. With a vector sampled from a distribution $D_{Z^m,\sigma}$, the conditional distribution is given as $D_{\Lambda_q^y(B),\Lambda}$. The n-dimensional lattice Λ is defined by $\Lambda = L(B) = \{Bx \colon x \in Z^m\}$. When $m \geq 2$ the basis of a lattice are concatenated by unimodular transformations $U \in Zn \times n$. In other words, lattices generated by a basis $B \in Z_q^n$ and another basis $\grave{B} \in Z_q^n$ are related by the expression $\grave{B} = BU$.

2.1.4 Discrete Gaussian

A Gaussian function $\rho_{\sigma,c}(x) = e^{\frac{-\pi\|x-c\|^2}{\sigma^2}}$ is defined with a target vector $c \in Z_q^n$ and Gaussian parameter $\sigma \in Z_q^n$ where $n \geq 0$ and $\Lambda \in Z_q^m$ is an n-dimensional lattice of points generated by a basis. The discrete Gaussian distribution is defined as follows

$$\forall z \in Z_q^n, D_{\Lambda,\sigma,c}(z) = \frac{\rho_{\sigma,c}(z)}{\rho_{\sigma,c}(\Lambda(B))} \tag{2.5}$$

$$\rho_{\sigma,c}(\Lambda(B) = \sum_{c\in\Lambda} \rho_{\sigma,c}(z) \tag{2.6}$$

$$\frac{\rho_{\sigma,c}(Bz)}{\rho_{\sigma,c}(\Lambda)} = \frac{e^{\frac{-\|Bz-c\|^2}{2\sigma^2}}}{\sum_{z\in Z^n} e^{\frac{-\|Bz-c\|^2}{2\sigma^2}}} \tag{2.7}$$

Definition 1. *There exists a smoothing parameter $\eta_\epsilon(\Lambda)$ which is a lower bound on the Gaussian parameter such that $\rho_{1\setminus\sigma}(\Lambda^* \setminus \{0\}) \leq \epsilon$ for $\epsilon > 0$ and $\Lambda \in Z_q^n$.*

To construct a Gaussian distribution $D_{\Lambda+c,\sqrt{\Sigma}}$ by a support with coset $\Lambda + c$ then the distribution becomes

$$D_{\Lambda+c,\sqrt{\Sigma}}(x) = \frac{\rho_{\sqrt{\Sigma}}(x)}{\rho_{\sqrt{\Sigma}}(\Lambda + c)} \propto \rho_{\sqrt{\Sigma}}(x) \tag{2.8}$$

where $\Sigma = B^T B$. For any orthogonal matrix $L, \Sigma = (LB)^T(LB)$ and for any set of integers Z^n, the Gaussian function $\rho_{B,c}(Z) = \sum_{x\in\Lambda+c} \rho_{B,c}(x)$. The cumulative distribution function $(cdf)_{\sigma,c}(x) = \sum_{i=-\infty}^{x} D_{\sigma,c}(i)$.

2.1.5 Learning with error

The Learning with Error assumption processes a noisy pseudo-random sampler Θ_s that generates a random uniform vector $s \in Z_q^n$ from a distribution $\chi \in \Psi_\alpha$ where $q \geq 2$. The LWE instance also employs a random sampler $\Theta_\$$ that generates a random fresh pair (w_i, v_i) where $v_i = (w_i^T s + y_i) \in Z_q^n \times Z_q$ are uniformly distributed over the domain with a noisy vector $y_i \leftarrow \Psi_\alpha$. The LWE problem outputs a toss of coin $l \in \{0, 1\}$ as a function of $poly(n)$ number of instances which is generated from a noisy pseudo-random sampler or a random sampler. The hardness of the LWE problem is summarized in theorem 1[58]

Theorem 1. *If there exists a quantum algorithm that solves the LWE instance then there exists an efficient quantum algorithm that finds the solution to the shortest vector problem to within $\Theta(n \setminus \alpha)$ in the worst case scenario given n, α with $q \geq 2$ and bounded on input such that $\alpha q \geq \omega(\sqrt{\log m})$.*

The theorem is an improvement on [59]proposition which states that solving the shortest vector problem on a lattice point is equivalent to solving a problem on a lattice of n dimension to the value $poly(n)$ factors and improved to $\Theta(n)$ factors. There is also the ring variant where the lattice is over a set of real numbers.

2.1.6 Basis reduction

The LLL algorithm [7] approximates a solution to the shortest vector problem by using the Gram-Schmidt orthogonalized equivalent of the basis of the lattice and non-negative integers. Through a series of swaps with introduction of appropriates indices $k \in \{1, 2, \dots n+1\}$, the Euclidean distance of the real number is reduced to less than half of its length $|\mu_{i,j}| \leq \frac{1}{2}$ for $1 \leq j < i \leq n$. Also the condition $|\tilde{b_i} + \mu_{i,i-1}\tilde{b_{i-1}}|^2 \geq \frac{3}{4}|b_{i-1}|^2$ for $1 < i \leq n$ is met if $k = n+1$. The swapping process depends on the size of k. If $k \leq n$ then $b_k, \mu_{k,j}, \mu_{k,k-1}$ is swapped with $b_k - sb_{k-1}, \mu_{k,j} - s\mu_{k-1,j}, \mu_{k,k-1} - s$ respectively, where s is a non-negative integer. If $k \geq 2$ then $\tilde{b_{k-1}}, \mu_{k,k-1}, \mu_{k-1,j}, \mu k, j, \mu_{i,k-1}\mu_{i,k}$ is swapped with $\tilde{b_k} + \mu_{k,k-1}\tilde{b_{k-1}}$ for $j < k-1$. The projection of the basis on the orthogonal complement is given by $\tilde{d_{k-1}} = \tilde{b_k} + \mu_{k,k-1}\tilde{b_{k-1}}$. The square of the Euclidean distance of the Gram-Schmidt orthogonalization is less than the square of the Euclidean distance of the basis of the lattice for $1 \leq i \leq n$. This is because the square of the Euclidean distance of the projection is less than three quarter of the complement of the Gram-Schmidt orthogonalization. If $k \geq 2$ and some aspects of the Gram-Schmidt orthogonalization is changed, the overall sum of the determinant of the lattice changes.

2.1.7 Basis Delegation

This method is employed to generate lattice basis for cryptosystem construction[60]. They can be categorized into SampleBasisLeft and SampleBasisRight. In literature, a short basis is used but the definition of the basis delegation method can be extended by replacing it with the Gram-Schmidt orthogonalization
SampleBasisLeft$(B, U, \tilde{B}, \sigma) \rightarrow T$: This probabilistic method outputs a basis $T \in Z_q^{n \times m}$ when given a reduced basis $\tilde{B}$ of $\Lambda_q^{\perp}(B)$, a random matrix $U \in Z_q^{n \times m}$, a Gaussian parameter σ and basis $B \in Z_q^{n \times m}$ with full rank on the condition that the Gaussian parameter $\sigma > \|\tilde{B}\|.\omega(\sqrt{\log m})$ where n, m and q are integers and $m > 2n$.
SampleBasisRight$(B, U, \tilde{B}, \sigma) \rightarrow T$: This randomized algorithm outputs a basis $T \in Z_q^{n \times m}$ when given a reduced basis $\tilde{B}$ of $\Lambda_q^{\perp}(B)$, a random matrix $U \in Z_q^{n \times m}$, a

Gaussian parameter σ and basis $B \in Z_q^{n \times m}$ of $\Lambda_q(B)$ with full rank on the condition that the Gaussian parameter $\sigma > \|\tilde{B}\|.\sqrt{m}.\omega(\sqrt{\log m})$ where $m > n$ and $q > 2$.

2.1.8 Least square Error based Basis construction

By finding the square of the error given as $||s - Bh||^2 = ||D_m\beta||^2 = ||H_m^*||^2$ using the solution given as $h = B^{'}s = (B^H S)^{-1}S^H x = [R_n\beta_{n+1}^T + Z_m\alpha_n]$, the minimum squared error function using the least square method can be estimated. Since the basis vectors of a lattice occupy orthogonal positions in the Euclidean space, an orthogonal matrix whose Hermitian transpose H_m^* performs a faster computation on the basis is constructed. It can also be shown that an arbitrary augmented matrix [61] is equivalent to the household transformation of the basis vectors[48] which reinforces its geometric properties.

$$\begin{bmatrix} R_n & \vdots & \alpha_n \\ \cdots & \cdots & \cdots \\ \beta_{n+1}^T & \vdots & \Theta_n \end{bmatrix} = \begin{bmatrix} R_n & \vdots & Z_m \\ \cdots & \cdots & \cdots \\ 0 & \vdots & g_m \end{bmatrix} = H_m^* \begin{bmatrix} \omega B_{m-1} \\ \cdots \\ \Gamma_m^T \end{bmatrix}$$

$$H_m^* = \begin{bmatrix} R_n \\ \cdots \\ \beta_{n+1}^T \\ \end{bmatrix}$$

$$\mathrm{B} = \begin{bmatrix} \alpha_n \\ \cdots \\ \Theta_n \end{bmatrix} \quad (9)$$

where ${\beta_{n+1}^T}^1 = [\ l^T : 0^T] \ \in Z^n$, $H_m^* \in Z^{m \times m}$ with entries modulo q and ω is a weighting factor.

[1] padding n-m zeros to l

$$\begin{bmatrix} \omega B_{m-1} \\ \dots \\ \Gamma_m^T \end{bmatrix} = \tag{10}$$

$$\begin{bmatrix} \omega b_{1,1}(0) & \omega b_{2,2}(-1) & \dots & -\omega b(-m+1) \\ \omega b_{1,1}(1) & \omega b_{2,2}(0) & \ddots & -\omega b(-m+2) \\ \vdots & \vdots & \omega b_{n,n}(m-1) & \vdots \\ {}_m(n-1) & {}_m(2) & {}_m(N) & \Lambda_q(n-m) \end{bmatrix}$$

Expanding H_m^* further a basis that has a triangular structure is given as

$$\begin{bmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,N} \\ & b_{2,2} & \dots & b_{2,N} \\ & & \ddots & \\ & & & b_{n,n} \end{bmatrix} \tag{11}$$

From (11), it can be seen that the basis is in a Hermite normal form. The triangu-

Algorithm 1 Household Triangularization[48]

for i=1:N **do**
- $\alpha_i = \sqrt{(\omega b_{i,i})^2 + (\Pi_{i-1}\Psi_m)^2}$
- $\sigma_i = \alpha_i + |\omega b_{i,i}|$
- $\rho_i = 1 - \frac{\Pi_{i-1}\Psi_m)^2}{\sigma_i}$
- $eta_i = \omega b_{i,i} + max\{|Z^n|, |\Lambda_q|\}\alpha_i$
- $\Psi_i^{'} = \frac{\Psi_i}{\sigma_i}$
- $\mu_i^{'} = \frac{\Pi_{i-1}\Psi_m}{\sigma_i}$ $\quad \omega b_{i,i} = max\{|Z^n|, |\Lambda_q|\}\alpha_i$
- $\omega^{'} = -\omega b + eta_i^{'}(\eta_i \omega s + \Pi_{i-1}\Psi\Lambda_q)$
- $\delta = \frac{-\omega b_{i,i}\Pi_{i-1}^2\Psi_m}{\alpha_i}$
- $\Lambda_q^{(i+1)} = \Lambda_q^{(i)} + \mu_i^{'}(\eta_i \omega b + \Pi_{i-1}\Psi_m\Lambda_q^{(i)})$

larization algorithm requires n square roots, $16n$ multiplications and $9n$ additions. The norm of the basis S is expressed as $||S|| \leq \sqrt{2n-1}||\alpha_n||$

2.1.9 Lattice Based Construction

In this section, the methods employed to encrypt data using lattice based method is enumerated. The methods constitutes the steps used in identity based encryption were a master secret key is generated from identities[62]. The process of generating the master secret key is shown in Algorithm 2. A list UL is initialized to null to enable the storage of identities in the key generation phase. The key generation phase is shown in Algorithm 3, were the discrete Gaussian sampler is used

to generate short vectors that is used as secret key for encryption. The encryption process is shown in Algorithm 4 were a noise vectors are added to the subset sum to generate the ciphertext. Finally, the message is decrypted as shown in Algorithm 5.

Algorithm 2 Setup

Require: security parameter λ maximal number of users, $N = poly(\lambda)$, identity space $ID = Z_q^n \leftarrow \{0,1\}$ Parameters q, n, m, σ, α, dimension $ID = \{1\} \times Z_q^{l-1}$

Ensure: (PP, MSK, UL)

1: set lattice parameters q, n, m, σ, α where $q \geq 2, m > 0, m > 3n, \sigma = L.\omega(\sqrt{})\alpha\sqrt{m} \geq \eta_{epsilon}(\Lambda^{\perp})$
2: Compute T_B, $(B, T_B) \in Z_q^{n \times m} \times Z_q \leftarrow TrapGen$(q, n, m, λ) where $\|T_B\| < L$
3: Set hash functions $\{0,1\}^l \leftarrow H_h^* : Z_q^{n \times m}$
4: Compute $C_{(R)}$ and $D_{(R)}$
5: select $s \in_R Z_q^n$ where $B = Rotf(s)$
6: set $MSK = T_B \quad PP = (H_h^*, B, C_{(R)}, D_{(R)}, s) \quad UL := \oslash$
7: **if** $i = i + 1$ **then**
8: Return (PP, MSK, UL)

Algorithm 3 Key Generation

Require: $ID \in \{0,1\}$, public parameters

Ensure: $SK_{id} = \{e_l\}_{l \in [d]}$

1: set the bound on the tailcut parameter Psi_ϵ and store ID in UL
2: Choose random vector $\{C\}_{l \in [d]} \in Z_q^n$
3: Compute $D_{(R)} \in Z_q^{n \times m}$ from $D_{(R)} \leftarrow B + H_h^*(ID)G$
4: **for** $l \in [d]$ **do**
5: $e_l \leftarrow$ SampleDiscGaus $(B, \tilde{B}, \sigma, c, r_i^*, D_{(R)})$ where $r_i^* = \|b_i^*\|$ is the norm
6: Return SK_{id}

Algorithm 4 Encryption

Require: $M \in \{0,1\}$, public parameters (B, D, H_h^*), identity $ID = (1, ID_1, \ldots ID_l)$

Ensure: $CT_{ID} = (C_1, C_2) \in Z_q \times Z_q^{3m}$

1: select $R_i \leftarrow \{-1, 1\}^{l \times m}$
2: choose $s \leftarrow Z_q^n$
3: set $C_{(ID)} = BR_{ID} - H_h^* D_{(R)}$
4: Compute $R_{ID} = \sum_{i=1}^{l} b_i Ri$ where b_i is a linearly independent vector of the basis
5: select $x, y \leftarrow \Psi_\alpha$ where $\Psi_\alpha \leftarrow Z_q^m$ and x, y noise vectors
6: Compute $C_1 = B^T s + y + M[\ \frac{q}{2}] \in Z_q$
7: Compute $C_2 = C_{(ID)}^T w_i + \begin{bmatrix} x \\ R_{ID}^T x \end{bmatrix} \in Z_q^{3m}$
8: Return CT_{ID} =0

Algorithm 5 Decryption

Require: $CT_{ID} \in Z_q \times Z_q^{3m}, e \sim D_{Z,\sqrt{\Sigma},t_i}, s \leftarrow Z_q^n$

Ensure: $\{0,1\} \leftarrow M^{'} \in Z_q$

1: Compute $M^{'} \leftarrow C_1 - s^T e \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$
2: **for** $i = i + 1$ **do**
3: **if** $|M^{'} - [\ \frac{q}{2}]\ | < [\ \frac{q}{5}]$ **then**
4:
5: Return 1
6: **else**
7:
8: Return 0

2.1.10 Correctness

Let e be the vector sampled from the discrete Gaussian distribution, let $\begin{Bmatrix} y \leftarrow \Psi_\alpha \\ x \leftarrow \Psi_\alpha \end{Bmatrix}$. Let $s \in Z_q^n$ be the secret vector from the noisy pseudorandom sampler Θ_s and a fresh vector w_i from random sampler $\Theta_\$$, then the secret key becomes $w_i^T s \in Z_q^n$,. Consider the ciphertext $(C_1, C_2) = (B^T w_i^T s + y + M[\ \frac{q}{2}] \in Z_q, C_{(ID)}^T w_i + \begin{bmatrix} x \\ R_{ID}^T x \end{bmatrix} \in Z_q^{3m})$ and the decryption algorithm becomes $M^{'} \leftarrow C_1 - s^T e \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$ which is equivalent to $B^T w_i^T s(1 - s^T e) + y(1 - s^T e) + M[\ \frac{q}{2}]\ (1 - s^T e - C_{ID}^T w_i s^T e + x^T s^T e - R_{ID}^T s^T ex)$. By lemma $\|e\| \leq \sigma\sqrt{m}$ and $Be = H_1(ID)$, therefore $M^{'} \leftarrow x^T s^T e - R_{ID}^T s^T ex + M[\ \frac{q}{2}\ ]$. Consequently, the algorithm outputs M, if $x^T s^T e - R_{ID}^T s^T ex$ is close to $\frac{q}{5}$ from the origin of the n-dimensional space modulo q and outputs 0 if

otherwise. From Lemma $x^T e$ is distributed with zero mean and standard deviation $\sigma\sqrt{m}\alpha q\omega(\sqrt{\log n})$ with high probability and the norm of the random uniform R_{ID} is bounded by $\Theta(\sqrt{m})$. Therefore, resulting to $x^T s^T e = \|x^T s^T e\| \leq \sigma\sqrt{m}\alpha q\omega(\sqrt{\log n})$ with high probability and $R_{ID} = \|R_{ID}\| \leq \Theta(\sqrt{m})$ with high probability. This follows that $\|R_{ID}^T s^T ex\| \leq \Theta\sqrt{m}$ and $\sigma\sqrt{m}.\alpha q\omega(\sqrt{\log n}) \leq \Theta(\sigma\alpha m\omega)$. The error term is bounded by $\|x^T s^T e\| + \|R_{ID}^T s^T ex\| \leq \alpha\omega\sqrt{m}q\omega(\sqrt{\log n}) + \Theta(\omega\sigma\alpha m)$ where αq is the ciphertext noise addition.

2.1.11 Security Analysis

The security of the proposed construction would be analyzed as a set of hybrid interactions between the adversary and the challenger in what constitutes indistinguishability against chosen ciphertext attack(IND-CPA).

Hybrid 1: This is the original IND-CPA game

Hybrid 2: Let $(id^*) \cap Q = \oslash$ be the challenge id set. This is the same with Hybrid 1 but the key generation algorithm is modified given matrices β_{n-1}^T and $B_{(R)}$ chosen at random, the Simulator selects $R_i^* \in \{-1, 1\}^l$ which is sent and generates $C_{(R)}$ and $D_{(R)}$ as follows $R_n^* \beta_{n-1}^T - H_m(B_{(R)}) = C_{(R)}$ and $R_n^* \beta_{n-1}^T + H_{m^*} B_{(R)} = D_{(R)}$ where $C_{(R)}$ and $D_{(R)}$ are uniformly distributed in $Z_q^{n \times m}$. The adversary A select $t = \lceil log_q(2|Q|)^{id} \rceil$ and $q \geq 2|Q|^i d + |T|$ where $|Q|^{id}$ is the maximal number of queries for $A_{(R)}$ and $B_{(R)}$ and $|T|$ size of time space and passes it to the simulator C. C aborts if $Pr[\frac{1}{\omega_{ij}}] = H_h{}^*$. By lemma, distributions $(A_{(R)}, A_{(R)} R_i^*, (R_i^*)^T y)$ is statistically close to $(A_{(R)}, C_{(R)}| \ldots |C_{(R)}|D_{(R)}| \ldots |D_{(R)})$ consequently $(A_{(R)}, C_{(R)}| \ldots |C_{(R)}|D_{(R)}| \ldots |D_{(R)},$ $(R_i^*)^T y))$, thus $Pr[W_0] = Pr[W_1]$

Hybrid 3: This is the same with the previous game except in A's view the vector $v_o + \sum_{i=1}^n id_i(z_i bi)$ is seen as a private key for identities $id = (id_i, \ldots, id_n) \in \{0, 1\}^n$. The simulator C then chooses $n + l + k$ public key vectors v_i according to distribution $D_{s(n+l+k),\sigma}$ and generate linearly independent vectors c_i by computing $Bv_i(\mod q) = c_i$. Distribution of $v_o + \sum_{i=1}^n id_i(z_i bi)$ is statistically close to $D_{s\sqrt{m}\sum_{i=1}^n \frac{id_i}{(n+l+k),\sigma}}$ where $\frac{s}{(n+l+k)} = \omega \geq \eta \in (\Lambda_q^{\perp}(B)$ where B is a public parameter matrix.

Hybrid 4: For a secret randomly chosen vector $s \in Z_q^n$, a LWE problem is defined by sampling an oracle Θ which can be sampled from either random distribution $\Theta_{\$}$ or a noisy pseudo random distribution Θ_n. The adversary A is used to construct a simulator S to solve the problem. The simulator S requests from the oracle Θ unassigned vectors $(u_i, v_i) \in Z_q^n \times Z_q^n$ for m_i instances where $i = 0, \ldots, m + 1$. It

constructs the model as follows, transform $C_{(R)} \in Z_q^{n\times m}$ from m_i instances by fixing the ith columns with linearly independent vectors $v_i \in Z_q^n$. It also transform $D_{(R)} \in Z_q^{n\times m}$ from m_i instances by fixing the ith columns with linearly independent vectors $u_i \in Z_q^n$. Then $C_{(R)}$ and $D_{(R)}$ is constructed for $i \in [m]$ as in the previous game using ${H_h}^*$ and $R_i^* \in \{-1,1\}^l$. Also, the rth row of the LWE instance is assigned to $v_o \in Z_q^n$. Thereafter, send $(A, \{C_1, \ldots, C_n, \{D_1, \ldots D_n\}, T_B, v_o\}$ to the adversary. A responds with $id^* = (b_1^*, \ldots, b_l^* \in \{-1,1\}^l$ and the message bit $M^* \in \{0,1\}$. The game is aborted in the previous game if $id = id^*$. In the challenge phase, the challenger queries the linearly independent vectors $v_o, \ldots, v_n \in Z_q^n$ from the LWE instance and gets $v^* = \begin{bmatrix} v_1 \\ v_n \end{bmatrix} \in Z_q^n$ and $v^{**} = \begin{bmatrix} v_1 \\ v_{m+1} \end{bmatrix} \in Z_q^n$. It employs $R_i^* \in \{-1,1\}^l$ to compute $R_{id}^* \longleftarrow \sum_{i=1}^n bi^* R_i^* \in Z_q^{n\times 2m}$, CT^* as $C_i^* = \begin{bmatrix} v^* \\ (R_{id}^*)^T v^* \end{bmatrix} \in Z_q^{2n}$ and $C_2^* = \begin{bmatrix} v^{**} \\ (R_{id}^*)^T v^{**} \end{bmatrix} \in Z_q^{2n}$. Adversary A makes queries $b \in \{0,1\}$, then S returns $CT^* = \{C_i^*, C_2^*\}$ as response if $b = 0$ and returns uniform values $CT^* \leftarrow Z_q^{2n+k}$ if $b = 1$. When the oracle Θ is noisy in the distribution Θ_s, v^* is uniform in Z_q^n and v^{**} is uniform in Z_q^n and (C_i^*, C_2^*) is uniform in $Z_q^{2n\times 2n}$ thus $F_{id^*} = (A \setminus AR^* \setminus B)$. By distribution $v^* = A^T s + y$ and $v^{**} = B^T s + e$ for some random vectors $y, e \in Z_q^n$ distributed in $D_{\Lambda,s,\omega}$. Finally, resulting to $C_1^* = \begin{bmatrix} A^T s + y \\ (R_{id}^*)^T A^T s + (R_i^*) y \end{bmatrix} = \begin{bmatrix} A^T s + y \\ (AR_i^*)^T s + (R_i^*)^T y \end{bmatrix} = (F_{id^*})^T s + \begin{bmatrix} y \\ (R_i^*)^T y \end{bmatrix}$ $C_2^* = \begin{bmatrix} B^T s + e \\ (R_{id}^*)^T B^T s + (R_i^*) e \end{bmatrix} = \begin{bmatrix} B^T s + e \\ (BR_i^*)^T s + (R_i^*)^T e \end{bmatrix} = (F_{id^*})^T s + \begin{bmatrix} e \\ (R_i^*)^T y \end{bmatrix}$.
Hybrid 5: Given the parameters $A_{(R)}, C_{(R)}, D_{(R)}$ and $pk^* = (e^*, A_2^*)$ and $id^* = (id_i^*, \ldots id_l^*) \neq ID_i$. S picks equal message bit $(M_0, M_1) \in \{0,1\}$ and generate $C_0^*, C_1^* \in Z_q^{2m} \times Z_q$ as follows $u^T s + e^* + [\ \frac{q}{2}]_1$ and $A_2^T e^* + e^* + [\ \frac{q}{2}]_0$ where $u = v_0 + \sum_{i=1}^l id_i v_i$ and e^* distributed according to $D_{u_l, A_0/G/R_{ID},\omega}$ where T_G is a trapdoor for the basis G if $(H(ID)) - H(ID^*))G$ and $ID \neq ID^*$. If A outputs a valid ciphertext $C_0, C1 = C_0^*, C_1^*$, then S aborts.

2.1.12 Parameters

For the trapdoor generation algorithm, the dimension of the lattice should be bounded by $m \geq (1 + 4\epsilon)n$ where $q \geq 2$. The norm of the reduced basis $\tilde{B}$ is bounded by $\|\tilde{B}\| \leq \Theta\sqrt{m}(n)$ with a noise error vector bounded by $\alpha < \frac{2\sqrt{m}}{q}$ where $q \geq \beta\omega(\sqrt{m})$ and the constant $\beta = 3m(1+\epsilon)\sigma\omega(\sqrt{\log n})$. This proves the worst case hardness of the construction. The sampling algorithm works perfectly if the

Gaussian parameter is set $\sigma \geq \|\tilde{B}\|\sqrt{m(1+\epsilon)}\log q\omega(\sqrt{\log n})$. The error term is less than $\frac{q}{5}$, if the error rate is bounded as $\alpha \leq \omega\sqrt{n}m^{\frac{3}{2}}\omega(\sqrt{\log n}) + \sigma \leq \frac{q}{5}$. The error vector can increase the probability of decryption failure if it is too large. If the lattice dimension n is assigned as the security parameter λ that is $n = poly(\lambda)$, then the column of the lattice m is expressed as $m = n^{1+\alpha}$ where $n^{\alpha} \geq logq$. The leftover hash lemma satisfies $m \geq (n+1)\log q + \omega(\log n)$. From the analysis, the dimension of the basis plays a role in the storage cost and efficiency of the construction at the expense of security. It is very important to keep it in focus when designing the algorithm and choosing appropriate parameters that would guarantee performance.

Table 2.1: Storage cost with lattice dimension m, security parameter n and prime q

Public key Parameters	**Secret Keys**	**Ciphertext**	**Security**	
$(3mn + n + n^2N/2)log_2q$	$(2m)$	$(2m+1)log_2q$	128	[63]
$3n$	n	$3n$	128	[64]
$2n\log^2 q(2n+1)$	$2n\log^2 q$	$2n\log^2 +$	128	[65]
$(mn + (l+1)n)$	m^2	$3m\log^2 q$	128	[56]
mn	$mn\log 2q$	$mn\log(12\sigma)$	128	[66]
mn	mn	$3m$	128	Ours

2.2 Codes

Cryptosystems have been designed that employ the Hamming metric, which specifies the number of vector coordinates that distinguishes one vector from another. A digital signature generation algorithm whose security is dependent on decoding codes with rank metric called RankSign [67] have also been developed. An encryption scheme using Gabidulin codes based on the security assumption of the hardness of the Rank syndrome decoding problem and Decisional Rank Syndrome decoding problem was developed with the scheme achieving 2140 bits of security at a smaller public key. Furthermore, it was reported that the cryptosystem has small public key size than public key sizes from Moderate Parity Density Check codes [68]. The effect of algebraic attacks on the parameters of the RankSign algorithm in polynomial times and its vulnerability is due to the fact that the Augmented Low Rank Parity Check codes have low Hamming weight. The algebraic attacks tend to expose the trapdoor function employed in constructing the cryptosystem. A code based encryption scheme with rank metric and whose selective security proof is in the random oracle model was constructed. Its security lies on the hardness

of solving the Rank Syndrome Decoding (RSD) problem, Rank Support Learning (RSL) problem and the Augmented Low Rank Parity Check Code (LRPC+) problem. The RankSign Algorithm was employed to generate the trapdoor function and the binary tree was used for key updates with the complexity of the key updates increasing logarithmically[69]. By finding low weight codewords distinguishable from a random code, the cryptosystem can be broken [67].

2.2.1 Notation

Table 3.1 gives a summary of some notations used in this section.

Table 2.2: Notations

Symbols	Meaning
q	power of prime
F_q	finite field of q elements
F_{q^m}	extension field of degree m
F_q^n	vector spaces of dimension n over F_q
A	$n \times m$ matrix
a	vector
$G_q(n)$	set of subspaces belonging to F_q^n
$\ker(A)$	kernel of matrix A
ω	matrix multiplication complexity exponent
E	Subspace code E(minimum entropy subset of $G_q(m)$)
F	Subspace code F(minimum entropy subset of $G_q(m)$)
$\mathrm{ds}(E, F)$	subspace distance between E and F
$E \oplus F$	Smallest subspace
$\mathrm{rk}(A)$	rank of A over F_q
$\langle A \rangle$	F_q span of A

2.2.2 Syndrome Decoding Problem

The breaking of the public key in code based cryptography under the rank metric depend on the solution to the syndrome decoding problem in the rank metric which is defined as

Definition 2. *Find a vector x_i with rank weight w, and an integer which represent the ith column of a vector that lifts a codeword from X to $X^{'}$, if $H^T x = s$ where $s \in_R F_{q^m}^{n-k}$ is a syndrome and H is a parity check matrix over F_{q^m}.*

The definition can be extended to the low rank codeword problem as $\text{rank}(A - C) = w$ where $A \in F_q^{m \times n}$ is a matrix, C is a linear matrix code and w is the rank weight. Furthermore, the computational syndrome decoding problem which relates the negligible advantage of probabilistic polynomial time adversary to find the vector x given the parameters (H, s, w) is defined as

Definition 3 (Computational Syndrome decoding problem)**.** *The advantage of a probabilistic polynomial time adversary A to find a vector $x \in F_{q^m}$ such that $H^T x = s$ after making adaptive queries q_τ to obtain a syndrome s where $H \in F_q^{n \times m}$ and w is the rank weight is with negligible advantage ϵ.*

It is very important that the probability of an algorithm that would solve the rank syndrome decoding problem is to be bounded by 1. Two other security assumptions employed in the rank syndrome problem are stated as follows.

Definition 4 (Decisional Rank Syndrome decoding problem)**.** *A probabilistic polynomial time adversary has negligible advantage to compute G a generator matrix where $G \in F_{q^m}^{k \times n}$ and $x^{'}$ from $MG + x$ given a message bit $M \in F_{q^m}^k$ and a vector $x \in F_{q^m}^n$ with row weight w. In other words, $Adv_{m,n,k}^{drsd} = Pr[\ A(G, mG \oplus x) = (G, x^{'}] \leq \epsilon$ where $x^{'} \in F_{q^m}^n$.*

The definition of the Decisional Rank Syndrome decoding problem can be extended as

Definition 5. *A probabilistic polynomial time algorithm has negligible advantage to compute $H^{'}$ from an augmented LRPC code $H^{'}$ with a homogeneous parity check matrix $H \in F_q^{n-k \times n}$ and $x^{'}$ from $MP^T s^{'} + x$ given a message bit $M \in F_{q^m}^k$, a syndrome $s^{'} \in F_{q^m}^n$ a vector $x \in F_{q^m}^n$ and square invertible matrix $P \in F_{q^m}^{n-t \times n}$. In other words, $(Pr[\ A(H, MP^T s^{'} + x) = (H^{'}, x^{'})]\ \leq \epsilon$*

In the chosen plaintext attack, the adversary makes adaptive key generation queries and the challenger responds by running a key generation algorithm with input of a security parameter and responds with a master public key. The adversary chooses two equal message bits $|M_1| = |M_o|$ and sends to the challenger who responds

by running the encryption algorithm to generate ciphertexts. The Adversary chooses random bits $b \in \{0,1\}$ and outputs a bit $b^{'} \in \{0,1\}$. The adversary wins if $b = b^{'}$.

2.2.3 Code based construction

Setup: Let $H^{'}$ be a parity check matrix of Augmented LRPC code with a parity check matrix H of weight $d \geq 4w + 1$ and P is a square and invertible matrix with weight $\frac{w}{2}$. Given a security parameter λ, a function $K\colon \{0,1\}^{*} \mapsto F_{q^m}^{k+t}$ is selected. A generator matrix $G \in F_{q^m}^{k \times n}$ with syndrome s is constructed. This generator matrix is decodable if $d \leq 2w$ where w is the weight of the vector $x \in F_q^n$. Then a random vector $u \leftarrow F_{q^m}^n$ is generated with rank $||u|| = n$. The master public key is $T = (P, s)$ and public parameters; (RP^T, G, u) where $R \in F_{q^m}^{(n-t)\times t}$.

Key Generation: Compute the square and invertible matrix P as $P \leftarrow K(ID)$ where $P \in F_q^{k+t}$. Then choose $u \in_R F_{q^m}^{n+l}$ and compute a syndrome $s = H^{'}x^T$ where $x \in F_q^{n^{'}+t}$. Finally, $sA = P - x$ is computed which follows $s^{'} = \frac{P-x}{A}$ and then return $SK_{ID} = \{s_i^{'}\}$.

Encryption Compute the square and invertible matrix P by $P \leftarrow K(ID)$ with the message bit $m \in F_{q^m}^{n^{'}} - k^{'}$. Then compute the ciphertext $C_1 = m(H^{'}G + P^T s^{'}) + x$ and $C_2 = C_1\left(\frac{RP^T}{x}\right) + C_1\left(\frac{I}{s^{'}G}m\right)$ and return $CT = (C_1, C_2)$ where $R \in F_{q^m}^{t\times n}$.

Decryption

$$(s^{'}|-1)(\frac{C_2}{C_1}) = s^{'}C_2 - C_1 = s\left(C_1(\frac{RP^T}{x}) + C_1(\frac{I}{s^{'}G})m\right) - C_1$$

$$= s^{'}C_1(\frac{RP^T}{x}) + s^{'}C_1(\frac{I}{s^{'}G})m - m(H^{'}G + P^T s^{'}) + x$$

$$= s^{'}C_1(\frac{RP^T}{x}) + C_1(\frac{I}{G})m - mH^{'}G + MP^T s^{'} + x$$

Since $H^{'}$ can be expressed as its constituent identity matrix and G is the generator matrix of H then it follows that $s^{'}C_1(\frac{RP^T}{x}) + mP^T s^{'} + x$. Since $d \leq 2w$ then $mP^T s^{'} + x$ can be solved with an efficient algorithm to recover the message.

2.2.4 Security Analysis

Theorem 2. *The Advantage of a probabilistic polynomial time adversary to win the IND-CPA game after making adaptive q_H and q_{KG} queries to the random oracle is $\epsilon \leq (q_H + q_{KG})\epsilon_{SDP} + \epsilon_{ffvp}$ where $\epsilon, \epsilon_{sdp}, \epsilon_{csdp}$ are negligible functions.*

Proof. The game is the same with the original IND-CPA game except the adversary A makes adaptive queries to the oracle $K : \{0,1\} \mapsto F_q^k$ for the challenge set id_j^*.

A makes adaptive query for the secret key and the challenge responds with the parameter (P, s, u) and runs the key generation algorithm $KG(PP, msk, id) \rightarrow SK_{ID}$ which sends SK_{ID} to A

In the challenge phase A makes more queries on the challenge id not queried and outputs 2 equal messages M_0 and M_1. If $id^* = id_i$ the challenger C aborts the game even though the guess bit β is hidden from A otherwise A returns a bit c^* and wins if $c^* = c$. The bit c is generated at the random toss of a coin by the challenger and used to compute the ciphertext CT which is sent to A. This game is indistinguishable from the original game by the hardness of the syndrome decoding problem that is

$$|Pr[\ G_1]\ - Pr[\ G_0]\ | \leq \epsilon_{sdp}(q_H + q_{KG})$$

After an adaptive queries by the challenger on the identity id^*, A generates a challenger ciphertext $CT^* = (C_1^*, C_2^*)$ from a random matrix R.

Since x^* is sampled from a discrete Gaussian distribution, the second game is indistinguishable from the first game by the hardness of the computational syndrome decoding problem that is

$$|Pr[\ G_2]\ - Pr[\ G_1]\ | \leq \epsilon_{csdp}$$

□

2.2.5 Information set decoding from linearized polynomials

The solutions to the syndrome decoding problem is reviewed in the rank metric. The problem is modeled and solved using linearized polynomials that form a set of quadratic equations.

2.2.5.1 Hauteville-Tillich

In the Hauteville-Tillich algorithm, the linear matrix code C is lifted to $C^{''}$ by multiplying two $n \times m$ invertible matrices P and Q to become $C^{''} = QCP$ with the same rank weight w and $C^{''} = \sum_{j=1}^{l} \beta_l$ where β_l is a basis of the subspace and x_i has zero entities with its ith entry equals to 1. Then choosing a random space V which is expressed as $V = \sum_{j=1}^{r} V_j$ and entries chosen uniformly for $j.\{l+1, \ldots, r\}$ and the basis chosen in this manner $V_j = 0 \forall j \in [\ 1'l]\ ; V_j = 0 \forall j \in [\ l+1, r]\ , V_i = 1 \forall i \in [\ 1, l]\ .$ Then the error vector x is expressed as $x_j = \sum j = 1^r \alpha_{ij} V_j$ for $i \in \{1, \ldots, m\}$. The number of variables that make up the equation becomes $(m-l)(r-l)+l(n-l) \forall i \in \{l+1, \ldots, m\}, \{l+1, \ldots, r\}$ and $j \in [\ 1, a]$ and the dimension of $C^{''}$ is given as

km. Therefore the number of equations is equivalent to $nm - km$. For a solution to occur, the number of equations should be $(n-k)m \geq (m-a)(r-a) + a(n-a)$. This results to $r = \lfloor \frac{m}{m-l}(n-k) + a\frac{m-n}{m-l} \rfloor$ and the resulting complexity becomes $O(n-k)^3 m^3 q^{(w-l)} \lfloor \frac{m}{m-l}(n-k) + l\frac{m-n}{m-l} \rfloor)$ [70].

2.2.5.2 Gaborit-Ruatta-Schrek

In the first phase, the linear matrix code C with parity check matrix H is lifted to $C^{'}$ by the expression $C^{'} = C + F_{q^m} x$ where $C^{'}$ contains the codeword x. To generate the linear matrix code, the parity check matrix H is computed from the solution of the rank decoding problem $Hx^T = s^T$, then the generator matrix G of the matrix code C is used to transform it into $C^{'}$. Since the dimension of the subspace F is α, then $x^{'}$ is of the form αx where $\alpha \in F_{q^m}$. Consequently, employing the support trapping approach, a codeword x such that $1 \in \text{support}(x) = E$ can be found, which follows that $1 \in F = E$ where E is a space and $F \supset \text{support}(x)$. When the codeword as a function of a basis is represented it becomes $x_j = \sum_{j=1}^{r} \alpha_{ij}\beta_j$ where β_i is a basis of the subspace and αij is the column space of the codeword x_j. Consequently, this results to $\dim\psi(F) = \dim F - 1$ where $\psi = \frac{V}{F_q}$ and V is a random subspace. Furthermore, expressing the subspace F as a function of its basis becomes $F = \sum_{l=1}^{n} \psi_l \beta l$. Therefore setting up the number of variables nr with m equations results to

$$\sum_{l=1}^{n}\sum_{j=1}^{r} \psi_l \alpha_{ij} H_l^{'} \beta_j +, \ldots, \sum_{l=1}^{n}\sum_{j=1}^{r} \psi_l \alpha_{ij} H_{n-k-l}^{'} \beta_j = 0 \tag{2.9}$$

Since $F \supset \text{supp}(x)$, there must be one solution and for this to hold, $m > n$ that is $m(n-k-l) \geq nr$ which results to $r = \lfloor \frac{m(n-k-l)}{n} \rfloor = m - \lceil \frac{m(k+1)}{n} \rceil$. Since $\dim\psi(F) = \dim F - 1$ and $\dim\psi(E) = \dim E - 1$, then the number of subspaces of dimensions w becomes $w-1$ and the number of subspace of dimension r becomes $r-1$, then the Gaussian probability becomes $P = \dfrac{\begin{bmatrix} w-1 \\ r-1 \end{bmatrix}_q}{\begin{bmatrix} w-1 \\ m-1 \end{bmatrix}_q} = q^{(w-1)(m-r)}$. The complexity of decoding x where $H^{'} x^{'T} = H\alpha x^T = s$ becomes $O((n-k)^3 m^3 q^{(w-1)\frac{(k+1)m}{n}}$[10].

2.2.5.3 Gaborit-Ruatta-Schrek Algorithm

Given a multivariate polynomial $F_m(f_1, \ldots, f_m) = \sum_{i=0}^{r} f_i x^{q^i}$ and another multivariate polynomial $F_y = (f_1, \ldots, f_m) = \sum_{i=0}^{r} f_i y^{q^i}$ all of degree r, based on

the linearity principle $F(\alpha x + \beta y) = \alpha F_x + \beta F_y \forall \alpha, \beta \in F_q$. if $\bar{x}$ and $\bar{y}$ are roots of the equation then $F_{\bar{x}} = F_{\bar{y}} = 0$. For a matrix code $C \in F_{q^m}$ with generator matrix $G = \sum_{i=1,j=n}^{k} g_{ij}$ and rank$(x) = r$ then $y = XG + x$ where $X \in C$ is a codeword. Given a random subspace $V = \sum j = 1^r V_j, X = \sum_{j=1}^{k} X_j$, there is $x_j = \sum_{j=1}^{r} \alpha_{ij} V_j$ that gives $k+r$ variables of the polynomial F_y. There exist a random subspace $V \subset E$ where $F_{\bar{y}} \in E$. Then for $\forall j, 1 \leq j \leq m, y = XG+x = F(y_j = F\left(\sum_{j=1}^{k} X_j g_{ij} + x_j\right) =$ 0 where $F(x_j) = \sum_{j=0}^{r} f_j x_j^{q^j}$. The unknowns are chosen in this manner; for $k + r$ unknowns $f_j x_j^{q^j} \forall i \in [\,1, k]$ and $j \in [\,0, r-1]$, for k unknowns $x_j^{q^r} \forall j \in [\,1, k]$ and for r unknowns $f_j \forall j \in [\,0, f_{r-1}]$(corresponding to scalar coordinate of y). Consequently the number of equations $m \geq (r + 1)(k + 1) - 1$, if $x_j = 0$, then X_j is decreased by one and then the variables are decreased by $(r + 1)$ terms [10].

2.2.6 LDPC codes

Given an edge e of the coordinate i of a subgraph shown in Figure 2.1 and Log likelihood ratio(LLR) of the channel message as it moves from the vertex of the LDPC coder to the LT coder at iteration t; $L_{c_j \to \nu_j}^{(i,t)}$, an edge e at the origin of the subgraph and its LLR at iteration $t + 1$; $L_{c_j \to \nu_j}^{(t+1)}$ and an edge e at coordinate i and coordinate j and its LLR of the channel message at iteration t as $L_{c_j \to \nu_j}^{(i,j,t)}$. The LLR can be related as follows [71]

$$L_{c_j \to \nu_j}^{(i,t)} = \frac{1}{2}(1 - \prod_{j=1}^{d-1}(1 - 2L_{c_j \to \nu_j}^{(i,j,t)}) L_{c_j \to \nu_j}^{(t+1)} = G(L_{c_1 \to \nu_1}, \ldots, L_{c_j \to \nu_j})$$

where $G = \frac{\prod_{i=1}^{n} \nu_j^{(t)}}{\prod_{i=1}^{n} \nu_j + \prod_{i=1}^{n} (1-\nu_j)}$ is the parent tree. For a subgraph with depth 2, the path along the edges for the vertex c_j of the LT code moving in a cyclic manner and terminating at the originating vertex c_j with girth is given by $2(2) + 2 = 6$. A trellis of $t \geq 2k$ can locate vertex c_j if and only if $2k \leq 10$ that is if the girth increases beyond 10.

If the bipartite graph with vertices V projects into the factor graph[72] G, then the block length is expressed as $n = \frac{|V^{'}|^2}{4} + |V^{'}|$ and the length of the information bits is expressed as $k = \frac{|V^{'}|^2}{4}$. The serial concatenation used in generating erasure codes is modelled using factor graphs which is actually a bipartite graph with edges and vertices. The degree of the individual vertices which is sampled from an output degree distribution $\Omega(x) = \sum^{i=1^d} \Omega_d x^d$ by randomly selecting an integer $m \in_r \{1, \ldots d\}$ should be less than or equal to the number of output symbols.

This implies that $\Omega(x)_{R\to\infty} = e^{\alpha(x-1)}$ where $\alpha = R$ and β is the average degree of the vertex that represents the source codewords given as $\beta = \sum_{i=1}^{d} d\Omega_d$. In other words, the average degree of the vertex that represent the output codewords. The edges that connects the vertices of the intermediate symbol and the output symbol is sampled from a distribution expressed as $\omega(x) = \sum_{i=1}^{d} \omega_d x^{d-1}$ where ω_d is the finite number of edges linking the output vertices. To optimize $\omega(x)$, density evolution is employed[73] . The posteriori rate of the code design[74][73] is expressed as the rate by which the degree of the edges are sampled from the distribution spanning from the intermediate vertex edges to the outer vertex and given as $R_{post} = \frac{1}{\alpha \sum \frac{\omega_d}{i}}$. The goal of the code design is to minimize R_{post} by employing the edge distribution of the output vertex. Consequently, the individual intermediate symbols is XORed to generate the codes which are transmitted over erasure channels [2]. The arbitrary

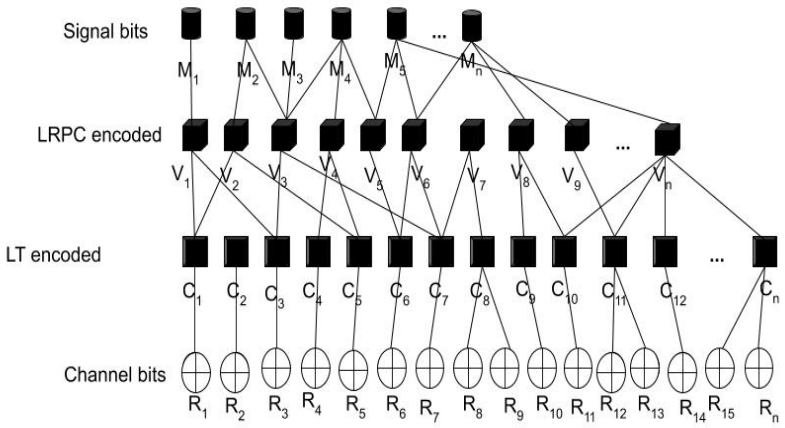

Figure 2.2: Non-planar bipartite structure from factor graph for erasure codes

number of symbols δ introduced to the source symbols during transmission generates a successful decoding τ^{succ} [74], which is analogous to the LLR of the intermediate vertices whose LLR $L_{\nu_j \to c_j} = 0$ and its probability density function is a function of the dirac delta function Δ_j[75]. The receiver overhead can be expressed as a function of a finite number of the output codewords $\xi = \frac{\delta}{k}$. The complexity of decoding

[2]BIAWGN channel modeled as $y_j = x_j + e_j$ where $x_j = (-1)^{-2c_j}$ and channel estimate $L(c_j) = 2y_j SNR$ where $SNR = \frac{1}{\sigma_e^2}$ and σ is the variance of the distribution $\mathcal{N}(-1, \sigma^2)$. The channel estimate can be extended to represent the channel log likelihood ratio as follows $L(c_j) = \log \frac{\rho_{c_j}(0/y_j)}{\rho_{c_j}(1/y_j)}$ where ρ_{c_j} is the probability density function as a function of the channel output symbol

is a function of τ^{succ} and inversely proportional to each other. The source generator matrix G relates the source vector $\vec{D}$ to the intermediate vector as $\vec{C} = G_1.\vec{D}$ where $G_1 \in F_{q^m}^{k \times n}$ while the intermediate generator matrix relates the intermediate vector to the output vector as $\vec{E} = G_2.\vec{C}$.

It can be seen that the distribution is Gaussian with mean $2SNR$. Employing the sum product approach[76], the number of output symbols k is increased at each iteration in order to recover the message. A problem arises when the symbols from the output vertex are exhausted, the transmission is aborted. In order to remedy the situation, a technique termed inactivation decoding[77] is employed which deactivates the remaining intermediate vertices from the selected decoding graphs which have not been processed on the condition that the maximum vertex in the generator matrix with the highest degree is selected and in the process the decoding operation continues. An increase in inactivation, leads to an increased received bits and the received bits is equivalent to deleting the intermediate vertex from the bipartite graph. In order to reduce the receiver's overhead, the symbols received must be used to generate linear systems whose generator matrix is full rank. In edge deletion decoding[78], the edges and vertices of the inactivated intermediate symbols are deleted to generate output vertices with a degree of one. Also, the output vertices are employed in the decoding graph which requires the degree distribution of the intermediate vertices from the unprocessed bipartite graph. The resulting output symbol degree distribution is given as [79]

$$\Omega_d(x) = \sum_{i=1}^{d} \Omega_{d,i} x^i = \sum_{i=1}^{d} (\sum_{\alpha_k=0}^{d-1} \Omega_{\phi+\alpha_k} \left(\phi + \alpha_k \alpha_k \right) (1-d)^{\phi} d^{\alpha_k}) x^i = n\Omega((1-d)x + d)$$

where ψ is the set of random degrees α_k. If the degree of the output vertex is one, then the average degree of the vertex of the source symbol goes to infinity. The LLR of the channel as the message moves from the source vertex ν_{LPDC} precoded by the source code LDPC to the intermediate vertex ν_{LT} at iteration i is given as

$$L_{\nu_j \to c_j} = 2tanh^{-1} \left(tanh \left(\frac{L(c_j^{(i)}}{2} \right) \prod_{\nu'_{LDPC} \neq \nu'_{LDPC}} tanh \left(\frac{L(c_j^{(i-1)}}{2} \right) \right) \right) \tag{2.10}$$

the LLR message of the first term can be expanded thus in relation to the mean of

the message[80]

$$E\left[tanh\left(\frac{L(c_j^{(i)}}{2}\right)\right]= \frac{\int_{-\infty}^{\infty} tanh\frac{u}{2}e^{-\frac{(u-\mu)^2}{4\mu}}du}{\sqrt{4\pi\mu}} \tag{2.11}$$

The LLR of the channel as the message moves from the vertex of the LT coder ν_{LT} to the vertex of the LPDC coder ν_{LDPC} at iteration i is given as [80]

$$L_{c_j\to\nu_j} = L(\nu_i) + \sum_{\nu_{LT}\neq\nu_{LT}^{'}} L_{\nu_j\to c_j} \tag{2.12}$$

where $L(c_j)^{(i)}$ and $L(c_j)^{(i-1)}$ are mutually and statistically independent. The mean of the message at ith iteration of the Belief propagation based message passing is given as [74]

$$\mu^{(i+1)} = \alpha\sum_{i=1}^{d}\omega_d L_{\nu_j\to c_j}(\mu^{(i)}) \tag{2.13}$$

In sum product algorithm, to approach Shannon capacity in binary erasure channels and converge the decoding error probability to zero, $\lim_{k\to\infty}\Omega(x) = 0$. This implies that the decoded block length must approach infinity.

Density Evolution is applied to the intermediate and outer symbol vertices in the bipartite graph to asymptotically model the density of messages as it moves from the edge distribution to the vertex distribution which determines the bound on the SNR. The BER is related to the code rate by the expression [75], $BER = 10\log_{10}(\frac{1}{2\sigma^2}).\frac{1}{R}$ and the output vertices are selected uniformly at random. In order to estimate marginals, $Pr(X = x)$, the statistical distance between the precode output vertex distribution and the marginal distribution should be minimized. In other words,

$$\sum_{i=1}\Omega_i x^{i-1} = \frac{1}{z_i}\sum t_i\Omega^{'}(x) \tag{2.14}$$

where t_i is a variable in the marginal distribution and $\Omega^{'}(x)$ is the partial belief state of the precode output vertex.

The following distributions is defined as follows;

$\Pi(x) = \sum_i \Pi_i x^i$; the intermediate vertex degree distribution

$\omega(x) = \sum_{i=1}\omega_i x^{i-1}$; output edge degree distribution

$\lambda(x) = \sum_{i=1}\lambda_i x^{i-1}$; precode intermediate nodes/intermediate edge degree distribution

$\Omega(x) = \sum_{i=1} \Omega_i x^{i-1}$; precode output vertex distribution. The distributions relate to each other as follows [73]

$$\omega_i = \frac{i\Omega_i}{\sum_{j=1}^{k} j\Omega_j}\lambda_i = \frac{i\Pi_i}{\sum_{j=1}^{n} j\Pi_j} \tag{2.15}$$

The Binomial distribution of the intermediate symbol encoding is given as

$$\Pi_1 = \binom{n}{i}\left(\frac{\Omega'(1)}{k'}\right)^i\left(1-\frac{\Omega'}{k'}\right)^{n-i} \tag{2.16}$$

After each iteration, the LLR of the message converges with the distribution of the intermediate symbol edges, that is $L_{\nu_{j_L T}} \rightarrow c_{j_L T} = \sum_{i=1} \Omega_i x^{i-1}$ in the bipartite graph. This ensures that the message is decoded, given that $BER \leq L_{\nu_{j_L T} \rightarrow c_{j_L T}}$. The whole essence of the design is to mitigate the scenario where the recorded bits are not enough to decode the message due to inactivation.

By employing density filtering[76] an update rule on the message of the vertex of the source coder can be given as follows

$$L_{\nu_{j_L T}} = \sum_{i=1}^{k} \frac{\Omega_i'}{\Omega_i'+1} + z[\sum_{d} \|1 - L_{c_{j_L T}}\|(L_{\nu_{j_L T}})^{(i-1)}]L_{\nu_{j_L PDC}} = \sum_{i=1}^{k} \frac{\lambda_i'}{d(\lambda_j'+1)^2} + z[\sum_{d} \|1 - L_{c_{j_L PDC}}\|(L_{\nu_{j_L T}})^{(i-1)}]$$

where z is the $E[\tan h(\frac{L(c_j^{(i)})}{2})]$

2.2.7 Finite design

By employing a function that maps the shaping matrix A to the Galois field of q elements, the parent tree can be coloured through $L_{c_j \rightarrow \nu_j}^{(t+1)}$. The function is defined as $\Gamma_k : A \rightarrow GF(q)(1 \leq k \leq d)$ and the colouring is defined as $e_j \in E \rightarrow \Gamma_k(\nu_j^{(t)}) \neq \Gamma(\nu_j^t)\forall j, s, k(O \leq j, s \leq m-1)$ and the width of the path the edge follows $\varpi_T(p)$ is expressed using a function [81]

$$\varpi_T(p) = \frac{\varpi_L}{\|\varpi\|}.\frac{1}{degG.N_p}.\prod_{q \in Prefix}\frac{1}{deg(q)-1} \tag{2.17}$$

where $L = \lceil \frac{|P|}{2} \rceil$ is the radius of the path. This bounds the number of vertices that have degree 0 assuming the subgraph is a null set. Let the absolute value of the eigenvalues of the parent tree be $\Delta = |\lambda_1| \geq |\lambda_2| \geq, \ldots, \geq |\lambda_d|$ and let the lower bound on the eigenvalues be $\leq 2\sqrt{\Delta - 1}$, resulting to [82]

$$N_{d=0} \leq \frac{4(\Delta-1)(1-b)n}{b\Delta^2} \leq \frac{4n}{b\Delta} \tag{2.18}$$

At the ith coordinate of n dimensional lattice, the coding gain can be related to the parent tree in the assumption that the subgraph can be projected to the subspace [83] which implies that

$$\frac{\phi_{\varpi_{T_P} \in F}(\Lambda)}{A_{\varpi_{T_P}}} = [\ \sqrt{\gamma(\Lambda)}\gamma(\Lambda)] \tag{2.19}$$

This implies that the mapping that characterizes the weight function of the path of the edges is a function of the coding gain and the area of the lattice subspace(resp. Babai nearest plane)[31]. It can be shown that the optimal weight coefficients of the code is equivalent to the capacity of the channel. By simplicity, denote the signal-to-noise ratio by S_{NR}. Then the optimal weight coefficients is defined [84] as follows

$$w_{i-1} = \sqrt{(1+S_{NRo})^i}\frac{S_{NRo}}{S_{NR}} \tag{2.20}$$

Squaring both sides, (26) becomes

$$\begin{aligned} w_{i-1}^{*2} = (1+S_{NRo})^{2i}\frac{S_{NRo}}{S_{NR}} &= (1+S_{NRo})(1+S_{NRo})\frac{S_{NRo}}{S_{NR}} = (1+S_{NRo}+S_{NRo}+S_{NRo}^2)\frac{S_{NRo}}{S_{NR}} \\ &= (1+2S_{NRo}+S_{NRo}^2)\frac{S_{NRo}}{S_{NR}} \\ &= (1+\mu_o+S_{NRo}^2)\frac{S_{NRo}}{S_{NR}} \end{aligned}$$

Let the signal-to-noise ratio be a function of an ideal[74] as follows $S_{NRo} = \texttt{info}(I \cap (\frac{\mu_o}{2\alpha}))$. Therefore, we have

$$w_{i-1}^{*2} = (1+\mu_o+S_{NRo}^2)\texttt{info}(I \cap (\frac{\mu_o}{2\alpha})) \tag{2.21}$$

In theory, the bound by fraction of output bits of degree 2 is given $\Omega_2 \geq \frac{\beta}{2\alpha\mu}$. then

$$w_{i-1}^{*2} = (1+\mu_o+S_{NRo}^2)\texttt{info}(I \cap \Omega_i) \tag{2.22}$$

Lemma 3. *Given an arbitrary integer $\eta \in R$ and as $n \to \infty$, the degree is scaled by a factor $1+\frac{1.08}{n}$*

Proof. For a constant expressed as $\delta = \sum_{n=3}^{\infty}(d_n - \frac{1}{n})\forall n, d_n = deg(\nu_n)$ there exist $\sum \delta_n = \log x + (\gamma + \delta - \frac{3}{2}) + O(\frac{1}{x})$. It can be seen that for values of $n \leq m$ where

$m \in [\ 0, 1]$ there is $\eta \in \left[\ \frac{(m-1)(m+1)}{(m-2)}, 1\right]$. This means that for an edge e of the parent tree, $w_n = d_n = \lfloor e(n-2)! - \frac{\eta}{n} \rfloor$ where $e_n = \sum_{k=0}^{n} \frac{1}{g_n}$ and g_n is the girth of the vertex ν. Given that $\sum_{k=1}^{\infty} \prod_{j=1}^{k} \frac{1}{n+j} < \frac{1}{n!} \sum_{k=1}^{\infty} \frac{1}{(n+k)^k}$ and for a depth of 2 of the subgraph, this reduces to $\sum_{k=1}^{\infty} \prod \frac{1}{n+2} < d_n < \frac{1}{n-2}$. From theorem [85] $\binom{n}{e}^{n-1} \sqrt{\frac{2\pi}{n}}(1 + \frac{1.08}{n}) < w_n < \binom{n}{e}^{n-1} \sqrt{\frac{2\pi}{n}}(1 + \frac{1.28}{n})$ since $n \to \infty$, the width path will converge to the degree. □

2.2.8 Information Set decoding from Optimization

To find the lifted codeword belonging to the transformed matrix code, an optimization based approach can be deciphered where the objective function $|x^{'}|_R = w$ is minimized such that $H^{'} x^{'T} = 0$ holds, where $H^{'}$ is the parity check matrix of a lifted code $C^{'}$. Given the error vector that lifts codeword X to $X^{'}$, the decoding problem can be redefined as; to find a codeword $x > 0$ that minimizes an optimization problem. In other words, the objective is to find a codeword $x \in F_q^n$ such that the function $\langle w, x \rangle$ is minimized with respect to $H \in F_q^{n \times m}$ such that $H^T x = s$ is satisfied. Such a codeword x is a solution to the problem. Assuming the syndrome is also lifted by an error vector e then there exist $s = He^T$, where $H \in F_q^{n \times m}$ is a parity check matrix and $e \in G$ where $G \in F_{q^m}^{k \times m}$ is a generator matrix. X a codeword such that $H^T x = He^T$ can be deduced which means that if $He^T = 0$, then, X can be found. It can also be seen that this satisfies the maximal decoding problem where $wt(xG + s) \leq w$ where w is the weight function of the edges..

Lemma 4. *Let E be a random subspace and F a subspace where $F \subset F_{q^m}$ then for any set of subspaces containing the restricted positions of the codeword x, there is*

$$supp(x) \subset F = \sum_{j=1}^{r} x_j \tag{2.23}$$

where $supp(x) = E$

Proof. The subspace E is equivalent to the set of coordinates of the codeword which is set in the subspace F over F_{q^m}. Let x_j be a solution to the problem $F(y_j = F(\sum_{j=1}^{k} x_j g_{ij} + x_j) = 0$. This is true by deduction. □

The following proposition can be expounded as, let x be such that the rank$(C) = r$ where $x \in C$, let β_j be the jth coordinate of the basis β_i of the space

$F \subset F_{q^m}$. Then the following holds $\dim\phi(\beta) = \dim F - \dim E = \dim F - \sum_{j=1}^{r} \beta_j = r$. Let $w_1, w_2^{'}$ be the dimensions of the space F and E where $w_1 = w_2^{-1}$, then by linearity $w_1\phi(w_2) = w_2^{-1}(F_{q^m}E) \leq \frac{q^m-1}{q-1}$.
The linear code can be transformed with the augmented LRPC code $H^{'}$ which is a function of the parity check matrix H of the linear matrix code. The codeword x is decoded optimally for a minimum $||x||$ where $||x||$ is the Euclidean norm of the codeword and $||y - x^{'}|| \leq \infty$. The objective function $||x||$ is minimized with respect to the augmented LRPC code $H^{'}$ such that $H^{'}y^T(\mod q)$ is satisfied where $y_0 - y = \mod q$ gives a lifted codeword $x^{'}$ and the $\text{supp}(x^{'}) \subset V$ is a random subspace such that $x - x^{'} = \mod q$. This implies that $x \mod q = x^{'}$. If $x_j = \sum_{j=1}^{r} \alpha_{ij} V_j \forall j \in [1, n]$ and $\alpha_{ij} \in F_{q^m}$ then $\deg F_q(\alpha(x_i), \ldots \alpha(x_n)(w - x_i) \leq 2r + k$ [*theorem*15] [86]]. This results to $(x_1, \ldots, V_1, \ldots V_{2r+k})$. For an injective isomorphism $\begin{matrix}\phi\colon F_q^m \mapsto F_q^n \\ x \mapsto \gamma_n\end{matrix}$ where $\gamma_n \in E$ and $\text{supp}(\gamma_n) \subset V$ then $\dim(\phi) = \dim(F) - \dim(V) = \gamma - (2r + k) = k - r$. A random choice of the basis is given as follows $V_j \in \{1, \ldots, n\}$ and $V_j \in \{n + 1, \ldots, n\}$ where $\sum_{i=1}^{n} \alpha_i V(y_i - x_i) = 0$. The syndrome decoding problem becomes
$H^{'}x^T = \sum_{l=1}^{n} \sum_{j=1}^{k} \alpha_{ij} H_l^{'} V_j = 0$. This implies that

$$\text{Prob}(V \subset F) = \frac{\begin{bmatrix} r - (n+k) \\ w - n \end{bmatrix}_q}{\begin{bmatrix} m - n \\ w - n \end{bmatrix}_q} = q^{(w-n)(m-r)} \tag{2.24}$$

For a solution to be found, $r - (n+k)(w-n) + m(2r+k) = m(n-k)$ which results in a complexity of $O(n+k)^3 m^3 q^{(w-n)\left\lfloor \left(\frac{m}{k-1}\right) - 1 \right\rfloor}$. From the injective isomorphism $\phi\colon F_q^m \mapsto F_q^n$, employing multivariate polynomial perspective, it follows from lemma 5,

Lemma 5. *Let $X \in C$ generated by a Generator matrix with parity check matrix H and let $h_1, \ldots, h_n$ be the coordinates of H, then $X = \sum_{i=1}^{n} \phi X_j h_{ij} = J_x$ where J_x is a function of x where $x = x_i\beta_i, \ldots x_n\beta$ and $\phi\colon F_q \mapsto F_{q^m}$.*

Proof. if there exist a solution x_j to the multivariate polynomial then $\sum_{i=1}^{n} \phi F(x_j) = 0$ satisfied resulting to

$$\sum_{i=1}^{n} \alpha_i F(y_j - x_j) = \sum_{i=1}^{n} \alpha_i (F(y_j h_{ij}) - J_x(h_{ij})$$

$$= F(y_j \sum_{i=1}^{n} \phi_i h_{ij}) - J(\sum_{i=1}^{n} \phi_i h_{ij}) = F(y_j(x_i) - J(x_i)$$

$$= F(y_j)(x_i) - F(x_i) + J(x_i).$$

By linearity principle $F(y_j)(x_i) + F(\alpha x_i + \beta x_i)$. Since $F(x_i) = 0$ which reduces to $F(\alpha x_i + \beta x_i) = \alpha F_{x_i} + \beta J_{x_i}$ □

2.3 Multivariate polynomials

Multivariate polynomial and its irreducibility has been studied in literature [87]. This irreducible property defines its security against quantum cryptanalysis thereby making it a good candidate for designing Post quantum cryptosystems. The security of Multivariate cryptography is based on the hardness of solving random nonlinear multivariate quadratic equations over finite fields. An instance of these equations is the Hidden Field equations(HFE). Due to its weakness, variants of the HFE have been developed which employ vinegar equations [88]. It has been recommended that the finite field should be small, a finite field of characteristic 2 GF(2) is appropriate. However, this comes at an expense in public key size which arises from large design parameters. It has been discovered that it is easier to hide the structure of an injective mapping into a large codomain than to hide the structure of a bijective map into a codomain of the same size as the domain [40]. Key-recovery based on eliminating redundant polynomials were carried out with success in breaking a multivariate scheme of 80-bit security level, albeit at increased computational overhead. [89].

In the hybrid approach [24] involving exhaustive search and Gröbner basis, small k variables are chosen and they are evaluated based on values chosen at random, in order to reduce the system to a zero dimensional set of multivariate equations. The XL algorithm [90] reduces the system to a linear system by multiplying the monomials up to an algebraic degree $D - 2$. If the degree is large, then a probable solution can be found. For a degree that is less than $D - 2$ the system is reduced by the Crossbred algorithm [91] to a linear system in which the first k variables has a degree D. The k variables generate k equations in which $deg_k > 1$ are removed and those whose $deg_k \leq 1$ are stored. To solve the resulting

linear system, the remaining $n-k$ variables is solved recursively using Gaussian elimination. This is prior to the $n-k$ variables being inserted using the Fast Evaluation approach. In the Unbalanced vinegar approach [88] approach, the variables are divided into fixed variables and the variables for determination. If the resulting linear system has most coefficients that is equal to zero then the system is solvable. For k columns of the linear system, m equations of the form $\sum_{i=1}^{m} \beta ij y_i^2 + y_i L_{i,j}(y_{m+1}, \ldots, y_{m+v}) + , \ldots + y_m L_{mj}(y_{m+1}, \ldots, y_{m+v}) + Q_j(y_{m+1}, \ldots, y_{m+v})$ after the first column of the linear system are fixed with random values. This method is inefficient when compared with Grover's search. To determine a bound, a set of solution such that the linear system is a member is usually searched and if a permutation defined as $\begin{cases} y_1, \ldots y_m \mapsto y_1, \ldots y_m \\ y_1, \ldots y_n \mapsto F \end{cases}$ is found, then the linear equations becomes zero [92]. Mapping the variables for insertion to variables for determination would lead to generation of quadratic terms. An instance of the Fast Evaluation technique is the Gray code enumeration where a function $f(x_1, \ldots, x_n)$ with coefficients b and $b^{'} \in F_2^n$ that differ at their ith columns are evaluated. For the expression $f(b) = f(b^{'}) + \frac{\delta f}{\delta x}(a^{'})$ to be satisfied, $O(1)$ machine instructions would be executed. The goal is to reduce $\frac{\delta f}{\delta x}(a^{'})$ to a constant. Gröbner basis computation generally has a tight exponential bound as the degree of the polynomial increases. In homotopy continuation computation, initial values of the root are not employed which guarantees a solution if the solution is close to the final approximated solution. However, they are advantageous in giving global solutions which are usually complex. Infact, 2^n initial values are usually considered. Also, it usually fails when the parameter t is varied between 0 and 1, in trying to move from the starting point to the solution. However, gradient descent proposed in Chapter 5 as a sparse solver is computational tractable. To the best of our knowledge, numerical methods like the gradient descent method proposed have not been studied extensively as an approach to cryptanalyze cryptosystem based on multivariate polynomials even when such polynomials have been reduced to its ideal. The basis of the polynomial is transformed to an unconstrained optimization problem that can be solved iteratively when a starting value close to the solution is appropriately guessed.

2.3.1 Medium Field multivariate equation

In this section, a core map would be recreated. The original core map was created from the approach by Wang [93] called Medium Field equations. In Wang's core map, the input variables are uniformly distributed and used to con-

struct the constituent matrices. However, this approach makes the system vulnerable to attacks. To solve this, the variables would be scrambled. This measure diffuses the core map as much as possible. The core map from the first instance is defined as $F\colon F_q^{12} \mapsto F_q^{16}$ where the map is made up of 12 variables that would form 16 output equations. Let $X_1, \ldots, X_{16} = \phi_1^{-1} \circ S(x_1, \ldots, x_{16r})$ and $Y_1, \ldots, Y_{22} = \phi \circ T(y_1, \ldots, y_{22r})$. Let $X_1, \ldots, X_{12} = \phi^{-1} \circ S(x_1, \ldots, x_{12t})$ and $Y_1, \ldots, Y_{16} = \phi \circ T^{-1}(y_1, \ldots, y_{16t})$ where $\phi^{-1}\colon F_q^{12} \mapsto F_q^{12t}$ and $\phi\colon F_q^{16} \mapsto F_q^{16t}$. A 2×2 matrices N_1, N_2 and N_3 is constructed as

$$N_1 = \begin{pmatrix} X_1 & X_5 \\ X_7 & X_{11} \end{pmatrix}, N_2 = \begin{pmatrix} X_2 & X_6 \\ X_8 & X_{12} \end{pmatrix}, N_3 = \begin{pmatrix} X_3 & X_4 \\ X_{10} & X_9 \end{pmatrix} \tag{2.25}$$

If the output matrices is constructed as $A_1 = N_1N_2$, $A_2 = N_2N_3$ and $A_3 = N_1N_3$, this results to

$$A_1 = \begin{pmatrix} Y_5 & Y_6 \\ Y_7 & Y_8 \end{pmatrix}, A_2 = \begin{pmatrix} Y_9 & Y_{10} \\ Y_{11} & Y_{12} \end{pmatrix}, A_3 = \begin{pmatrix} Y_{13} & X_{14} \\ X_{15} & X_{16} \end{pmatrix} \tag{2.26}$$

Expanding the matrices further, becomes

$$\begin{aligned} A_1 &= \begin{bmatrix} X_1X_2 + X_5X_8 & X_5X_6 + X_{11}X_{12} \\ X_7X_2 + X_7X_8 & X_{11}X_6 + X_{11}X_{12} \end{bmatrix} \\ A_2 &= \begin{bmatrix} X_2X_3 + X_6X_{10} & X_2X_4 + X_6X_9 \\ X_8X_3 + X_{12}X_{10} & X_8X_4 + X_{12}X_9 \end{bmatrix} \\ A_3 &= \begin{bmatrix} X_1X_3 + X_5X_{10} & X_1X_4 + X_5X_9 \\ X_7X_3 + X_7X_{10} & X_{11}X_4 + X_{11}X_9 \end{bmatrix} \end{aligned} \tag{2.27}$$

The variables Y_1, Y_2, Y_3 and Y_4 is now expressed as follows

$$\begin{aligned} Y_1 &= X_1 + X_5X_8 + X_6X_7 + Q_1 \\ Y_2 &= X_2 + X_9X_{12} + X_{10}X_{11} + Q_2 \\ Y_3 &= X_3 + X_1X_4 + X_2X_3 + Q_3 \\ Y_4 &= X_1X_5 + X_2X_5X_7 + X_3X_6X_7 + Q_4 \end{aligned}$$

where Q_1, Q_2, Q_3 and Q_4 are maps. The expression for the remaining variables from (15)can be found as follows

$$\begin{cases} Y_5 = X_1X_2 + X_5X_8 \\ Y_6 = X_5X_6 + X_{11}X_{12} \\ Y_7 = X_7X_2 + X_7X_8 \\ Y_8 = X_{11}X_6 + X_{11}X_{12} \\ Y_9 = X_2X_3 + X_6X_{10} \\ Y_{10} = X_2X_4 + X_6X_9 \\ Y_{11} = X_8X_3 + X_{12}X_{10} \\ Y_{12} = X_8X_4 + X_{12}X_9 \\ Y_{13} = X_1X_3 + X_5X_{10} \\ Y_{14} = X_1X_4 + X_5X_9 \\ Y_{15} = X_7X_3 + X_7X_{10} \\ Y_{16} = X_{11}X_4 + X_{11}X_9 \end{cases} \tag{2.28}$$

which satisfies the following condition

$$\begin{cases} \det(A_1) = \det(N_1).\det(N_2) \\ \det(A_1) = \det(N_1).\det(N_3) \\ \det(A_3) = \det(N_1).\det(N_3) \end{cases} \tag{2.29}$$

2.3.2 Zuang-Zi HFE(ZHFE)

In the ZHFE scheme [94] constructed from two HFEs, the affine transformation are constructed as $S\colon F_q^n \mapsto F_q^n$ and $T\colon F_q^{2n} \mapsto F_q^{2n}$ and the public key becomes $P_k = T \circ (\phi \times \phi) \circ (F_1, F_2) \circ \phi^{-1} \circ S\colon F_q^n \mapsto F_q^{2n}$ while the private key parameters are (Ψ, S, T, Ψ_D). To encrypt the plaintext $x \in F_q^n$ and generate a ciphertext $y \in F - q^{2n}$, $(y_1, \ldots, y_{2n}) = P_k(x_1, \ldots, x_n)$ is computed. If four polynomials were chosen at random $Q_{11}, Q_{12}, Q_{21}, Q_{22}$ where $Q_{kl}(Y) = \sum_{i=0}^{n-1} \upsilon_{kl,i} Y^{q^i}, (1 \leq k, l \leq 2)$, the map results to

$$\Psi(X, f_1, f_m) = X.[\ L_{11}(F_1) + L_{12}(F_2]\ + X^q[\ L_{21}(F_1) + L_{22}(F_2)]$$

The coefficients of f_1, f_m satisfy the conditions $deg\Psi(X, f_1(x_1), f_m(x_1 \leq D)$ where

D is the degree of regularity. Furthermore, the map parameterized at D is defined as

$$\Psi_D(X) = \Psi(X, F_1(X), F_2(X)) = \tag{2.30}$$

$$\sum_{0 \leq i \leq 1} \sum_{q^i + q^j + q^k \leq D} \alpha_{ijk}^{II} X^{q^i+q^j+q^k} + \sum_{q^i+q^j \leq D} \beta_{ij}^{II} X^{q^i+q^j} + \sum_{q^i \leq D} \gamma_i^{II} X^{q^i} \tag{2.31}$$

Further deduction, shows that it is hard to solve the following equations

$$\begin{aligned} f_1(X) &= Y_1 \\ f_m(X) &= Y_2 \end{aligned} \tag{2.32}$$

given Y_1 and $Y_2 \in F_q^n$. However to solve for X with complexity $O(nD^2 log_q D + D^3)$ then the following condition must be satisfied

$$\Psi_D(X) = \Psi(X, Y_1, Y_2) = 0 \tag{2.33}$$

In order to decrypt the plaintext, an inverse isomorphism is constructed as $(x_1, \ldots, x_{2n})$ $= T^{-1}(y)$. Computing $(Y_1, Y_2) = (\phi^{-1}(x_1, \ldots, x_n), \phi^{-1}(x_{n+1}, \ldots, x_{2n}))$ and substituting in Equation 2.33 results to

$$\Psi_D(X) = T^{-1}(y).(\phi^{-1}(x_1, \ldots, x_n), \phi^{-1}(x_{n+1}, \ldots, x_{2n})) \tag{2.34}$$

This can be solved with an efficient algorithm. Finally, $S^{-1}(\phi(X))$ is computed to check whether it is a solution to $y = P_k(x)$ and each solution derived for $x \in Z$ reduces the set of non- solutions.

2.3.3 Tame Maps

Tame multivariate schemes are constructed with Triangular Maps [95] of the form

$$\Gamma(x) = \begin{bmatrix} x_1 \\ x_2 + \gamma_1(x_1) \\ \vdots \\ x_n + \gamma_{n-1}(x_1, \ldots, x_{n-1}) \end{bmatrix} \tag{2.35}$$

where $\gamma_1, \ldots, \gamma_{n-1}$ are random polynomials. Since a triangular map is a bijective map then $\Gamma(x) = y$ has a solution which is gotten by mathematical induction. A triangular perturbation can be constructed by the addition of triangular maps as follows

$$\Gamma(x) = \Gamma(x_1, x_2) = \begin{pmatrix} x_{n+1} + \gamma_1(x_1) \\ x_{n+2} + \gamma_2(x_1, x_{n+1}) \\ \vdots \\ x_{n+s} + \gamma_s(x_1, \ldots, x_{n+s-1}) \end{pmatrix} \tag{2.36}$$

The core map becomes $F^{'}(x) = F(x_1) + B.\Gamma(x_1, x_2)$ where B is a randomly chosen matrix. The triangular perturbation is a surjective map which means that $\Gamma(x_1, x_2) = y$ has a solution which is computed through mathematical induction.

2.3.4 Construction

In this section the encryption process used in Multivariate polynomial cryptography is outlined. It involves three processes; Key generation, Encryption and Decryption. The key generation process is shown in Figure 2.2.

Key Generation Let F_q be a finite field with q elements and F_{q^d} be a degree d extension field over F_q. A vector space isomorphism is defined as $\phi\colon F_{q^d} \mapsto F_{q^d}$. Let k, l, v and n be integers and let $u = \{u_1, \ldots, u_k\}$. A degree n irreducible polynomial is chosen where $d > n$, $g(X) \in F[\,X]$ and $F_{q^d} = F_q[\,X]\,/\langle g(X)\rangle$. The public identity $id_i = H(P_i)$ is computed using a hash function H. Let $n = d + l + v$. Two invertible affine transformations $S\colon F_q^n \mapsto F_q^n$ and $T\colon F_q^n$ are chosen. An isomorphic map is defined as $\phi(u_1 + u_2 x +, \ldots, u_n x^{n-1}) = \sum_{i=1}^{n} x_i X^{i-1}$. An n variable quadratic polynomial public key is constructed as $P = T \circ \phi \circ F \circ \phi^{-1} \circ S(x_1, \ldots, x_n)$. A core map is defined as follows $P_k(x) = \sum_{i=0}^{n-1} \alpha_i^{(k)} x_i x_j + \sum_{i=0}^{n-2} \sum_{j=i+1}^{n-1} \beta_{ij}^{(k)} x_i x_j + \sum_{i=0}^{n-1} \gamma_i^{(k)} x_i + \delta^{(k)}$ where $\alpha_i^{(k)}, \beta_{ij}^{(k)}, \gamma_i^{(k)}, \delta^{(k)}$ are chosen at random over F_q. P_k is a HFE map which is

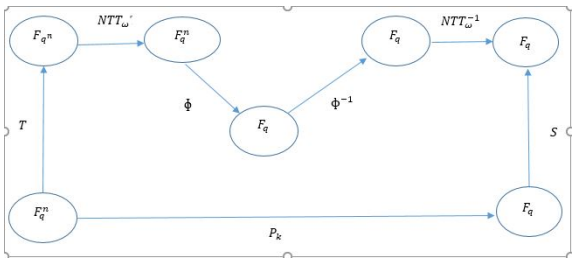

Figure 2.3: Key Generation

invertible. The private key are two invertible affine transformations S and T, scalars $\alpha_1, \ldots, \alpha_{2n}, \beta_1, \ldots, \beta_{2n}$ and an index i. Finally two mappings are defined; the projection onto the finite field of d coordinates $\pi\colon F_q^{k+l} \mapsto F_q^d$ and the linear embedding $\tau\colon F_q^{n+v} \mapsto F_q^d$.

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \rightarrow \begin{bmatrix} S_{1,1} & \ldots & S_{1,n} \\ \vdots & \ddots & \vdots \\ S_{n,1} & \ldots & S_{n,n} \end{bmatrix} \rightarrow \begin{pmatrix} \sum s_{1i}x_i \\ \vdots \\ \sum s_{ni}x_i \end{pmatrix} \rightarrow \tag{2.37}$$

$$\begin{bmatrix} F_{1,1} & \ldots & F_{1,n} \\ \vdots & \ddots & \vdots \\ F_{n,1} & \ldots & F_{n,n} \end{bmatrix} \rightarrow \begin{pmatrix} f_1(\sum x_1, \ldots, \sum x_i) \\ \vdots \\ f_m(\sum x_i, \ldots, \sum x_i) \end{pmatrix} \rightarrow$$

$$\begin{bmatrix} T_{1,1} & \ldots & T_{1,n} \\ \vdots & \ddots & \vdots \\ T_{n,1} & \ldots & T_{n,n} \end{bmatrix} \rightarrow \begin{pmatrix} \sum f_j(\sum x_1, \ldots, \sum x_i) \\ \vdots \\ \sum f_j(\sum x_i, \ldots, \sum x_i) \end{pmatrix}$$

Encryption Given a plaintext $m = (m_1, \ldots, m_n) \in F^n$, to generate a ciphertext $CT = (c_1, \ldots, c_n) \in F^m$, the core map $P_k(x_i)$ is used to mask the plaintext as $CT = P_k(m)$.

Decryption Given a ciphertext $CT = (C_1, \ldots, C_n)$ where $CT \in F^m$. By employing an inverse affine transformation S^{-1}, $y = (y_1, \ldots, y_n) = S^{-1}(CT)$ is computed. Also, $Y = \phi^{-1}(y_1, \ldots, y_n) \in F_q^n$ is computed. Consequently, an efficient algorithm is applied to compute solutions $X^{'}$ and $X^{''}$ for $F(X^{'}) = Y$ and $F(X^{''}) = Y^{'}$. For each solution $x = \phi(x_1, \ldots, x_n) \in F_q^n$ is computed. Finally, the ciphertext is decrypted by computing $(m_1, \ldots m_n) = T^{-1}(x)$.

2.3.5 Solvability using Quantum Approximations

An important process in cryptosystems especially quantum safe algorithms that involves irreducible polynomials is the effectiveness of finding solutions to the

polynomials in order to extract the original message. In this section, quantum approximation methods used in reducing the polynomial to a form where it can be solved iteratively to get the root of equations is discussed.

2.3.5.1 Grover's Search

The Grover's search is employed to generate M items of size N space in which the search problem is defined by a factor $h\colon \{0,1\}^n \rightarrow \{0,1\}$ such that $h(x) = 1$ [38]. In Reversible XL algorithm [96], the algorithm scans through $2^{n-1}q^f$ of possible values of the n variables of a polynomial f before the Reversible XL process is applied. If there is a solution to the linear system, the search method returns random solutions to choose from. The time complexity of the process is given as $q^{f/2+O(1)}$ quantum computations of the algorithm. This search method can be employed to quantize a function F and generate a variable $a_2 \in F_2^k$ such that $F^{cons}(a_2 = 1)$ where $a_2 = (a_{n-b+1}, \ldots, a_n)$ in the ClassicalBooleanSolve algorithm [97]. Search is faster than the brute force method, if the time required holds for $t_{ver} > t_{SAT}2^{-(n-g)}$. Consequently, $F^{cons}(a_2) = (f_1(x_1, \ldots, x_n, a_{n-b+1}, \ldots, a_n), \ldots,$ $f_m(x_1, \ldots, x_n, a_{n-b+1}, \ldots, a_n)$. This means that the search is done on the last b variables spreading in $a_{n-b+1}, \ldots a_n$ all mapping into the last b properties of a solution. By substituting $f_2^{(i)} + f_3^{(i}$ with these variables, $m.2^b$ linear equations can be generated in the pre-processing phase which defines the search procedure. It also evaluates the exhaustive search algorithm and improves classical complexity in the process. The total cost equivalent to the concatenation of the query complexity together with quantum circuit complexity and the diffusion step as given in Equation(2.38) [98]

$$\frac{\pi}{4}.\sqrt{\frac{N}{M}}.(Quantumcircuitcost) + (Diffusionstepcost) \tag{2.38}$$

Furthermore, the quantum circuit construction employs NOT gates and $(m+1)$-bit Toffoli gates. Also, the number of classical queries is given as $O(N/M)$. The Grover's search queries the quantum circuit and then carries out the diffusion process on the n-qubits. Thereby inverting the mean of the amplitude of the M items being searched in a space size of N. By iterating a single search procedure, the number of search processes is reduced to its square root [98]. The Quantum circuit is also called the quantum oracle or the blackbox.

2.3.5.2 Macaulay Matrices

To determine the consistency of Macaulay Matrices, sparse system solvers are employed due to sparse nature of the matrices. The algorithm outputs a cer-

tificate of inconsistency u when given a random matrix $A \in F_q^n$ and a vector b to give a solution to the linear system $Ax = b$. This algorithm performs matrix-vector multiplications with complexity of $O(n)$ and other operations with complexity of $O(n^2 \log n \log n \log n)$ [38]. For homogeneous polynomials, the columns of the Macaulay matrix are indexed by the coordinates of the polynomial by decreasing monomial order. For non-homogeneous polynomials, the columns are indexed by the terms of the function over the finite field of q element while the reverse is the case for the rows. This is expounded by the lemma [38].

Lemma 6. *Let* $F = (f_1, \ldots, f_m) \in F_q[\, x_1, \ldots, x_n]$ *and for* $1 < d < \frac{n}{2}$ *then* $C_{MAC} < \frac{1-x}{1-2x}\binom{n}{d}$, $r_{MAC} < m.\frac{x^2}{(1-2x)(1-x)}\binom{n}{d}$, $S_{MAC} < mn^2.\frac{x^2}{(1-2x)(1-x)}\binom{n}{d}$ *where* $x = \frac{d}{n}$ *and* C_{MAC} *is the number of columns of the Macaulay matrices,* r_{MAC} *is the number of rows of the Macaulay matrices and* S_{MAC} *is the number of non zero entries of the Macaulay matrices.*

The XL algorithm checks the consistency of the Macaulay matrices by trying to determine whether vector b is a non-zero linear combination of the rows of these Macaulay matrices up to degree 2 and these reduces to an equation that can be solved recursively with a fast root finding algorithm [99]. The random solve algorithm employs a quantum circuit to determine the consistency of Macaulay matrix. It produces a solution $x = -\sum_{i=1}^{n} \frac{f[\ i]}{f[\ o]\}A_r^i b_r}$ where $b_r = b + Aw$, w a random vector, $r = deg(f)$ and a Boolean Hilbert series $S_{m,n}$. This series gives a 1 if a solution is found and 0 if otherwise. The quantum circuit is constructed as

$$|a_{11}\rangle \ldots |a_{nn}\rangle |f[\ 0]\hat{-}1\rangle |f[\ 1]\ \ldots |f[\ n]\ \rangle |r\rangle w_1\rangle \ldots$$
$$|w_n\rangle |b_1\rangle \ldots |b_n\rangle |0\rangle \ldots |0\rangle \rightarrow |a_{11}\rangle \ldots |a_{nn}\rangle |f[\ 0]\ ^{-1}\rangle$$
$$f[\ 1]\ \rangle \ldots |f[\ n]\ \rangle |r\rangle |w_1\rangle \ldots |w_n|b_1\rangle \ldots |b_n\rangle |0\rangle \ldots |0$$
$$\rangle |b_1\rangle \ldots |b_1^{'}\rangle x_1\rangle \ldots |x_n\rangle |S_A\rangle$$

2.3.5.3 Quantum Circuit

Data stored in a quantum computer is approximated in qubits and the storage register is a function of the computational bases states $|x\rangle$ which is defined as [98]

$$|\Psi\rangle = \sum_{x \in \{0,1\}^n} \alpha_x |x\rangle$$

where $\alpha_x \in C$ is the amplitude. The probability of measuring the state is a function of the square of the amplitude such that

$$\sum_{x \in \{0,1\}^n} |\alpha_x|^2 = 1 \tag{2.39}$$

For a quantum circuit constructed from a Boolean function $h\colon \{0,1\} \to \{0,1\}$ with input n bits, output m bits and working memory w bits, the transformation is given as

$$h|x\rangle|y\rangle|\rangle\rangle^w \mapsto |x\rangle|y \oplus h(x)\rangle|g(x)\rangle$$

Chapter Summary

It can be concluded that the geometric properties of the lattice was reinforced using the householder transformation. Furthermore, the LLL approach which carries out unimodular transformation of the lattice basis was presented. Ideal lattices are useful in decreasing parameters due to their ring structure and secured due to irreducibility. Master public key sizes are exponential on the input of the security parameter. Triangularization can also be employed to construct lattice basis. Lattice basis are spanned by linearly independent vectors. The vectors are orthogonal to its rows. The vector from the subset sum creates a distribution that is statistical distant to a theoretical Gaussian distribution with negligible value. This error vector scrambles the lattice basis thereby obfuscating it from a quantum computer. The smoothing parameter forms a lower bound on the Gaussian parameter that defines the Gaussian distribution. This error vector is a basis for the Learning with error assumption and solving this instance is a prelude for a quantum computer to solve the shortest vector problem. By employing a series of swaps, the LLL algorithm reduces the Euclidean distance of the vectors to less than half of its norm. The swapping depends on the size of the index set. Basis delegation can be used to construct a lattice basis using a reduced basis and are categorized into SampleBasisLeft and SampleBasisRight. In the Key generation step, a discrete Gaussian sampler is used to generate short vectors for encryption. The security of a lattice construction is modelled as a set of interaction between the polynomial bounded adversary and a challenger under the IND-CPA attack. The game is aborted if both entities have equal identities. Furthermore, the error vector can increase the probability of decryption if is too large, so therefore there should be a range. Algebraic attack

on codes in the Hamming metric exposes the one-way function used in designing the system. By decoding low weight codewords given a parity check matrix and a syndrome, a rank, metric based cryptosystem can be broken. By multiplying two permutation matrices, a codeword can be lifted to be a function of a subspace and by solving overdetermined equations, a bound on the rank weight can be computed. The security of Multivariate schemes is based on the hardness of solving non linear equations over finite fields. By multiplying the monomials up to a certain degree defined by the degree of regularity, the system can be reduced to its ideal thereby making it a linear system. Examples of algorithms that reduce the system are the XL and Crossbred algorithm. A map can be bijective or injective and such maps are used in hiding the structure of the core map from interpolation by an adversary. For homogeneous polynomials, the columns of the Macaulay matrix are indexed to enable solvability of the linear systems containing ideals.

Chapter 3

Shortest vector problem: Solution using Dimensionality Mapping

3.1 Introduction

Lattice based cryptography has been the forefront of schemes that are designed to withstand attacks as a result of quantum computation [100]. The security of lattice cryptography is based on approximating vectors that lie in an n-dimensional space of discrete subgroups which has been proven hard in worst case scenarios as proposed by [59]. The public key size is a quadratic function of the security parameter where security is in the order of 512bits. This large key size makes it unpractical to employ lattice cryptography for energy constrained devices. This has led to the development of ideal lattices which are based on worst case scenarios which are equivalent to the hardness of average case instances over polynomial rings [101]. Many lattice based construction require a Gaussian sampling algorithm which is efficient and fast to sample lattice points given a basis and a vector. The vector is distributed with a centre c and a standard deviation that is close to the centre. It is possible to produce the vector without disclosing information about the basis. This reinforces its security proof in the midst of a quantum adversary. A solution to the shortest vector problem has been proposed using Klein's approach [30] which is a variant of Babai's nearest plane algorithm [31]. The complexity of Klein's approach was given as $O(n^3 log_2 B)$ while that of Babai's algorithm, is given as $O(n^4 log_2 B)$, where n is the lattice dimension, and B is the size of lattice basis made up of Gram-Schmidt vectors [102]. This complexity is as a result of not employing fast integer arithmetic but longer integer arithmetic which is based on

Gram-Schmidt orthogonalization. The drawback of floating point arithmetic is that floating point arithmetic require pre-computed tables which increase computation time[103]. The complexity of Gaussian sampling can also be attributed to the standard deviation of the distribution bounded by $\sigma \geq \omega(\sqrt{logn}).max_{1\leq i\leq n}||B||$ [27]. Markov chain based Monte Carlo algorithm has been proposed for sampling. This employs Markov chain constructed from Gibbs sampling algorithm [27] to generate samples from previous samples. Ducas suggested that to improve algorithmic complexity the Gram-Schmidt basis should pre-computed and stored before the sampling process which will lead to $O(n^2)$ operations [104]. It is usual that a system samples from a distribution statistically close to the theoretical Gaussian distribution to within 2100 by employing floating point operations that have precision of at least 100 bits[32]. Using standard (53 bit) double precision floating point numbers is efficient as compared to the use of multi-precision arithmetic numbers, but not up to 80bit or 100bit security levels.

Gaussian sampling is employed in lattice cryptography as an approach that solves the Shortest Vector problem (SVP). The shortest vector problem which states that finding a short vector in a given secret basis that has the same Euclidean norm as the shortest vector in a lattice is hard. This is because a basis is defined by the norm of its longest vector rather than the norm of the shortest vector. A solution to this problem is done by providing the best approximation to error vectors modulo lattice Z(Learning with errors) that lie in a uniform distribution when given a normal distribution with narrow width as shown in Figure 3.1.

This uniform distribution belongs to the Euclidean space R^n that is unbounded. A lattice is an n- dimensional space that is isomorphic to an Euclidean space and also its discrete subgroup. The basis lies in the lattice while the coefficient of the basis lies in the Euclidean space. The polynomial time adversary seeks the best approximation at different values of the standard deviation of the distribution. If the vectors are sampled accurately, a computational bounded adversary would be incapacitated to distinguish a simulated distribution of the Ciphertext from a uniform distribution. Consequently, intractability in distinguishing a Ciphertext from a plaintext when given a fixed message in a Chosen Ciphertext attack game. The Gaussian sampling algorithm involves an efficient reduction algorithm to reduce the lattice basis before a Monte Carlo based sampling method is employed to generate the vectors. The complexity of the lattice reduction algorithm depends

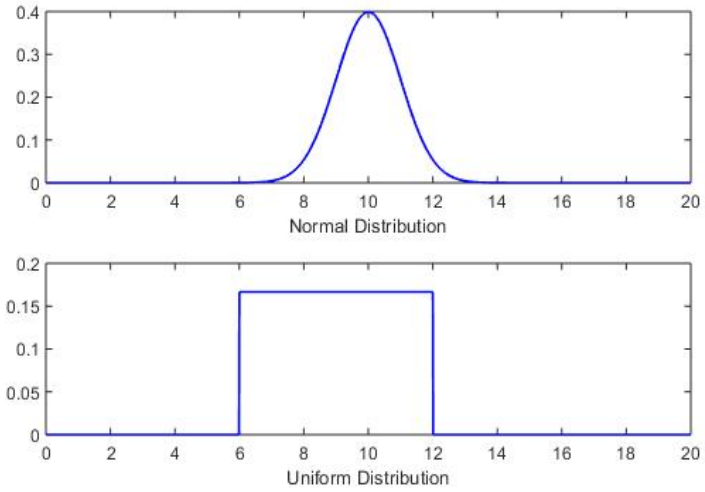

Figure 3.1: Approximation from normal to uniform with $\mu = 10, \sigma = 1$

on the magnitude of the uniform error vectors. The vectors have large Euclidean norms to makes it difficult for computationally bounded adversary to approximate an exponentially bounded uniform distribution on the input of a security parameter. This Euclidean norms are increased by raising the standard deviation of the distribution which transfers the probability density to the vectors which is discretized [105]. To limit decryption errors, the inner product term between the secret basis and the uniform error vectors would be bounded by a negligible n for some n. The angle of separation of the vectors is sampled from an interval $[\ -1/2, \ldots 0, \ldots 2/3]$ which satisfies condition of accurate approximation. This implies that vectors must be orthogonal to each other in space in order to converge exponentially to uniform distribution when given a normal distribution. In order to employ a 'cutting' mechanism to remove portions of the tail of the normal distribution, the Euclidean norm of the sampled vectors should be greater than the standard deviation of the distribution. If the distribution is uniform, the probability mass would be constant thereby distributing the norm evenly among the vectors. The tail of the distribution is the point where the discrete probability density magnitude is negligible while the continuous probability mass is infinite. It is intractable sampling at this point and is bounded by $|x| < \beta\sigma$. A smoothing parameter is parsed to remove this point and make cryptanalysis possible. The metrics of statistical indistinguishability described in this chapter are given as the statistical distance and the floating point precision defines the accuracy of sampling. Given the statistical distance, the polynomial

time algorithm cannot distinguish the simulated distribution from the perfect distribution by a probability bounded by the statistical distance. The precision defines the extent by which, given an adaptive adversary query, an approximation to the uniform distribution in polynomial time can be made. This is proportional to the inverse of the security parameter 2^{λ}. This implies that as the security increases in magnitude, the database of adversarial queries drops exponentially. In another variant of the Learning with errors(Ring-LWE), the coefficient of the error vectors are orthogonal to the canonical embedding by a factor. The significance of precision as a function of the security parameter has been studied in literature[106]. The lattice reduction reduces the basis to its Gram-Schmidt orthogonalization by using the LLL algorithm [7] which has the objective of constructing a basis with short Euclidean norms and orthonormal vectors. The LLL reduction has some drawbacks which has been mentioned in literature. They employ integers which have enormous precision and floating point arithmetic numbers which consume storage [107]. It also suffers from complexity and lack of constant time implementation [108] and finally difficulty in generating samples because the canonical embedding is made up of non-integer vectors [109]. In this chapter, a solution to this problem is proposed by employing Dimensionality mapping in which an embedding is constructed that transforms the basis from a point of high dimension to low dimension. Therefore, in the sampler, the dimension of the basis is exploited and not its orthogonality. This will offer tighter security proof and approximation.

3.1.1 Review of Works on Gaussian Sampling

Floating point arithmetic(FPA) approach to speed up Klein's algorithm involves producing quasi linear output which is a function of the size of input basis. This is usually small[110] when the ring is set at $R = X^n \pm 1$. This technique employs high and low precision FPA. The algorithm ran in $O(n^2)$ for optimized Klein's algorithm and $O(n)$ for Peikert's offline algorithm. The complexity was as a result of the higher precision[111]. An algorithm to compute the Gram-Schmidt orthogonalization basis using (n^2) arithmetic operations was developed. It was combined with Klein/GPV [112] sampling algorithm in which there is no pre-computation of the basis. However, the algorithm is slow because of the additional time in computing the basis and also sampling process. Furthermore, it was suggested that for the sampled distribution of vectors to be close statistically to the discrete Gaussian distribution, the precision of the Gram-Schmidt basis should be more than q bits.

However, it should not be too high to avoid numerical instability [113] . A practical Gaussian sampler that employed the cumulative distribution table algorithm was developed. In addition, an approach was proposed to reduce computation of the CDT by pre-computing for small number of centres in $O(n^3)$ which are not known prior to sampling. Furthermore, employing a lazy technique to compute the cumulative distribution function in double precision FPA when it is required. However, the proposed technique is not constant time making it impracticable[28][103]. However, by avoiding data dependent branching this problem can be solved. The output samples was expressed as a Boolean function of the input random bits. This is in order to generate them in constant time using bit slicing technique. The results were implemented using AVX vector software[114]. Metropolis-Hastings Klein algorithm has been proposed where the Markov chain is statistically close to the desired Gaussian distribution[27]. An efficient Hardware-based Cumulative distribution function inversion sampler was designed which samples with high precision and tail cut parameter. The statistical distance of the distribution was given as 2^{-90}. A small look up table was used to carry out faster operation. 9.44 random bits and 2.28 clock cycles were used to generate one sample [115]. A trapdoor sampling algorithm was implemented based on cyclotomic rings modulo prime using double precision floating point arithmetic [116]. The sampling procedure was categorized into an offline phase where the standard deviation and the centre of the distribution are not known and an online phase which was carried out in constant time[117].

In practice, sampling from discrete Gaussian distribution consumes more than 50% of the running time of a signature generation algorithm. Therefore, there is need to develop techniques to sample from discrete Gaussian distribution accurately and efficiently. Klein suggested a Gaussian parameter $\sigma \geq max_{1\leq i\leq n}\|\tilde{b}\|\omega(\sqrt{\log n})$ for his randomized version of Babai's nearest plane algorithm which samples from 1-dimensional Gaussian distribution while they proposed a Gaussian parameter $\sigma \leq max_{1\leq i\leq n}\|\tilde{b}\|\omega(\sqrt{\log n})$ [27] The target distribution approximates to a uniform random vector $x \in Z, 0 \leq x < 1$ by finding the shortest vector $i \in R, i \geq 0$ that is sampled from the interval $\frac{\sum_{i=-\infty}^{x} D_{\sigma,c}(i)}{\sum_{i=-\infty}^{x} D_{\sigma,c}(\infty)} \leq x < \frac{\sum_{i=-\infty}^{x} D_{\sigma,c}(i+1)}{\sum_{i=-\infty}^{x} D_{\sigma,c}(\infty)}$. The simulated distribution approximates to a uniform vector $j \in Z, 0 \leq j < 2^{\lambda}$ with high precision λ by finding the shortest vector $i \in R$ that is sampled from the interval $\frac{\sum_{i=-\infty}^{x} D_{\sigma,c}(i)}{\sum_{i=-\infty}^{x} D_{\sigma,c}(\infty)} \leq 2^{-\lambda} j < \frac{\sum_{i=-\infty}^{x} D_{\sigma,c}(i+1)}{\sum_{i=-\infty}^{x} D_{\sigma,c}(\infty)}$. Furthermore for accuracy, the vectors must be sampled at this interval $2^{-\lambda} j \leq \frac{\sum_{i=-\infty}^{x} D_{\sigma,c}(i)}{\sum_{i=-\infty}^{x} D_{\sigma,c}(\infty)} \leq 2^{-\lambda}(j+1)$[106] .

3.1.2 Contribution

This chapter proposes to reduce the basis using dimensionality mapping where the problem of reducing the norm of the basis vectors is considered as an optimization problem. In high dimensionality, approximation to the shortest vector problem is intractable and previous approaches on reduction only give a square root of the approximation[118]. Furthermore, there is reported loss of orthogonality during the reduction process[119] and at high dimensionality the deviation between the basis and an identity matrix increases[120]. This thesis solves the problem by constructing a low dimensional space of points $\tilde{B}$ and an affine transformation which is assumed to be rounded up integer belonging to a low dimensional set of integers Z^d. Then, an optimization problem is formulated which defines the shortest vector as a function of the Frobenius norm which would be the extended to the norm of the constituent vectors that make up the basis. The variants of Klein and Babai is extended by showing that the next iterate of the vectors would be an orthogonal function of the reduced center of the distribution and not an increment. Also, [26] convolution process is extended by reducing the covariance matrix using single value decomposition and then for a success probability greater than a negligible function ϵ, it is shown that the statistical distance between the theoretical and simulated distribution is close to 2ϵ.

3.2 Preliminaries

3.2.1 Gaussian Distribution

A Gaussian function is defined as $\rho_{\sigma,c}(x) = e^{\frac{-\pi\|x-c\|^2}{\sigma^2}}$ with a target vector $c \in Z_q^n$ and Gaussian parameter $\sigma \in Z_q^n$ where $n \geq 0$ and $\Lambda \in Z_q^m$ is an n-dimensional lattice of points generated by a linear combination of a basis. The discrete Gaussian distribution is defined as

$$\forall z \in Z_q^n, D_{\Lambda,\sigma,c}(z) = \frac{\rho_{\sigma,c}(z)}{\rho_{\sigma,c}(\Lambda(B))} \tag{3.1}$$

$$\rho_{\sigma,c}(\Lambda(B) = \sum_{c\in\Lambda} \rho_{\sigma,c}(z) \tag{3.2}$$

$$\frac{\rho_{\sigma,c}(Bz)}{\rho_{\sigma,c}(\Lambda)} = \frac{e^{\frac{-\|Bz-c\|^2}{2\sigma^2}}}{\sum_{z\in Z^n} e^{\frac{-\|Bz-c\|^2}{2\sigma^2}}} \tag{3.3}$$

where $Q = \sum_{z \in Z^n} e^{\frac{-\|Bz-c\|^2}{2\sigma^2}} = \sum_{Z^n \to \infty}^{\infty} e^{\frac{-\|Bz-c\|^2}{2\sigma^2}}$ is called the normalization factor[121].

Definition 6. *There exists a smoothing parameter $\eta_\epsilon(\Lambda)$ which is a lower bound on the Gaussian parameter such that $\rho_{1\setminus\sigma}(\Lambda^* \setminus \{0\}) \leq \epsilon$ for $\epsilon > 0$ and $\Lambda \in Z_q^n$.*

3.2.2 Linear Algebra

A symmetric matrix Σ can be decomposed to LD^2L^T where $L^TL = I$ and I is the identity matrix, D is the diagonal matrix of square roots of eigenvectors and L an orthogonal matrix. The column vectors of L which is orthogonal are the eigenvectors of the symmetric matrix Σ. The symmetric matrix Σ is positive definite if $x^T\Sigma x > 0$ for positive $x > 0$. Given two positive semidefinite matrices, Σ_1 and Σ_2, if $\Sigma \geq 0$ and $\Sigma_1 > \Sigma_2$ then $(\Sigma_1 - \Sigma_2) > 0$. If $\Sigma = BB^T$ then $B = \sqrt{\Sigma}$ for $\Sigma > 0$ where B is a singular matrix $B \in Z^{n \times m}$. When B is decomposed by single value process, it becomes LDM^T where L and M are orthogonal matrices and D is the diagonal matrix with positive entries. The expression $\rho_{\sqrt{\Sigma_1}}(x - c_1).\rho_{\sqrt{\Sigma_2}}(x - c_2) = \rho_{\sqrt{\Sigma}}(c_2 - c_1).\rho_{\sqrt{\Sigma_3}}(x - c_3)$ holds if $\Sigma_3^{-1}c_3 = \Sigma_1^{-1}c_1 + \Sigma_2^{-1}c_2$, $\Sigma_3^{-1} = \Sigma_1^{-1} + \Sigma_2^{-1} > 0$ and $\Sigma_0 = \Sigma_1 + \Sigma_2 > 0$ where x, c_1, c_2 are chosen uniformly from a distribution $D_{\sqrt{\Sigma}}$[116]. If z is sampled from a distribution with spherical structure then $x^T\Sigma x = x^T\Sigma_1 x + x^T\Sigma_2 x = \sqrt{\Sigma}.\frac{I}{2\pi}.\sqrt{\Sigma}^T = \frac{\Sigma}{2\pi}$ where $\frac{I}{2\pi}$ is the covariance of the vector z.

3.2.3 Gram-Schmidt Orthogonalization

In this section,the properties of the Gram Schmidt Orthogonalization is defined in Definition 7 and Lemma 7[113].

Definition 7. *Let $Span_l(B)$ be an l vector space with elements which are linearly independent and B is a basis with linearly independent vectors $B = \{b_1, \ldots, b_l\} \in Z^n$ is the root $\{\sum_{1 \leq i \leq l} x_i b_i, x_i \in X\}$. For any $x \in Z^n$, the projection of x over Λ is $proj(x, Span_x(B)) = x\tilde{B}(B\tilde{B})^{-1}B$ and for $y \in Z^m$, the projection of x over $Span_y(B) = proj(x, Span_y(B)) = \frac{\langle x,y \rangle Z}{\langle y,y \rangle Z}.y$.*

Lemma 7. *Let $B = \{b_1, \ldots, b_n\} \in Z^n$ be a basis for $l \in [\,1, n]$ and $\Lambda_l = Span_l\{B\}$, there is a reduced basis $\tilde{B} = \{\tilde{b_1}, \ldots \tilde{b_n}\} \in Z^n$ that can be defined with the properties as*

$$\forall l \in [\ 1, n]\, \tilde{b}_n = b_n - proj(b_n, \Lambda_{n-1}) \tag{3.4}$$

$$\forall l \in [\ 1, n]\, \tilde{b}_n = b_n - \sum_{i=1}^{n-1} \frac{\langle b_n, \tilde{n}\rangle l}{\langle \tilde{b}_n, \tilde{n}\rangle l} \tilde{b_n} \tag{3.5}$$

where $\forall l \in [\ 1, n]\, \tilde{b}_n \perp \Lambda_{l-1}$ *and* $(b_n - \tilde{b_n}) \in \Lambda_{l-1}$ *which follows from* $\tilde{\Lambda}_l = Span_l(\tilde{b_1}, \ldots \tilde{b_n})$ *to give* $\forall l \in [\ 1, n]\, \tilde{\Lambda}_l = \Lambda_l$.

3.2.4 Smoothing parameter

The smoothing or tailcut parameter $\tau \in \eta_\epsilon(\Lambda)$ for > 0 is the smallest value of the Gaussian parameter such that $\rho_{\frac{1}{\sigma},0}(\Lambda^* \setminus 0)$. This means a level of smoothness on the lattice Λ. Given computation $r > 4.72$ and $\sigma > 1$, the probability of the smoothing parameter in the Gaussian distribution becomes less than 2^{-100} [32]. The discrete Gaussian distribution with large radius that is equivalent to the tailcut parameter can generate a noise value over the lattice point to produce a uniform distribution. Due to the fact that the Gaussian parameter is transformed into a positive definite covariance matrix Σ, we have the expression $\rho_{\frac{1}{\sigma},0}(\Lambda^* \setminus 0) = \rho_{\sqrt{\Sigma}}(\Lambda^* \setminus 0) = \rho(\sqrt{\Sigma}\Lambda^* \setminus 0) \leq \epsilon$. Given infinite tail and high precision, a smoothing parameter that produces a uniform distribution can result in high computational operations. To solve this problem, a random bit is added after sampling which extracts the distribution for $x \geq 0$. In the process $\rho_{\sigma,0}$ is reduced by half. Lemma 8 defines the parameter further

Lemma 8. *For any* $\epsilon > 0, \sigma > \Psi_\epsilon(\Lambda)$ *and* $c \in Z^n$ *then the condition holds*

$$\frac{\rho_{\sigma,c}(\Lambda)}{\rho_{\sigma,0}(\Lambda)} \leq \frac{1-\epsilon}{1+\epsilon} = \rho_{\sigma,c}(\Lambda) \in [\ \frac{1-\epsilon}{1+\epsilon}, 1]\ .\rho_{\sigma,0}(\Lambda) \tag{3.6}$$

From the Gaussian sampling, z is sampled from a perfect distribution $Pr((\|z\|)c\sigma\sqrt{n}) < c^n e^{n/2(1-c^2}$ where n is the dimension of the lattice [122]. The result from this Lemma[123] is improved in Proposition 1

Proposition 1. *Let* $x_1 \leftarrow D_{Z,\sigma_1}, x_2 \leftarrow D_{Z,\sigma_2}$ *be sampled for some* σ_1, σ_2 *and let* $\sigma_3^{-2} = \sigma_1^{-2} + \sigma_2^{-2}$ *and* $\sigma^2 = \sigma_1^2 + \sigma_2^2$ *for any* $\epsilon \in (0, \frac{1}{2})$. *If* $\sigma_1 \geq \eta_\epsilon(Z)/\sqrt{2pi}$ *and* $\sigma_3 \geq \eta_\epsilon(KZ)/\sqrt{2pi}$ *then the distribution* P *of* $x_1 + x_2$ *satisfies*

$$D_{Kl}(P \| D_{Z,\sigma}) \leq 2(1 - (\frac{1+\epsilon}{1-\epsilon})^2)^2 \approx 32^2 \tag{3.7}$$

3.2.5 Statistical distance

Two random variables sampled from a proposed distribution $\chi \in \Psi_\alpha$ for a specified n dimensional space are statistically distant if the following holds

$$\Delta(t,z) = \frac{1}{2} \sum_{v,c \in \Psi_\alpha} |Pr[\ t = c] \ - Pr[\ z = v] \ | \tag{3.8}$$

The variable t defined over a distribution Ψ_α, its minimum entropy is given as $H_\infty(t) = -log(max_{c \in \Psi_\alpha} Pr[\ t = c]\ )$ where $max_{c \in \Psi_\alpha} Pr[\ t = c]$ defines the probability of detecting the random variable over the distribution Ψ_α. Wu defines the conditional minimum entropy as $\bar{H}_\infty(t,z) = -log(E_{v \in Z}[\ 2^{-H_\infty(t,z)}]\ )$ [124]. A probabilistic polynomial time adversary have a negligible advantage to distinguish between a variable sampled from a perfect distribution and a proposed distribution with a statistical distance bounded by $2^{\frac{-\Delta(t,z)}{2}}$ which is usually 2^{-90} to 2^{-128} [5].

3.2.6 Markov chain

When variables in a target distribution are not transformed into other variables with the same properties but different algebraic structure, a sampler can be used to build a Markov Chain which uses the previous state of a sample to generate the next state of the next sample. This sampler employs Markov Chain Monte Carlo algorithm for this purpose [27]. The term ergodicity is employed to denote the relationship between its rate of convergence or in other words how fast it decays the exponential that defines its statistical distance from the target to the proposed distribution and it is defined with the expression[27].

$$\lim_{T \to \infty} \|T^T(x,.) - D_\pi(.)\|_{TV} = 0 \tag{3.9}$$

where T is a transition matrix with vector x, $D_\pi(.)$ a stationary distribution and $\|\|_{TV}$ is the rate of decay in the exponential or asymptotic statistical distance. It is also uniform if this convergence is bound by $M(1-)^T$ for $0 < \delta < 1$ and $M < \infty$ where δ is the exponential decay coefficient.

3.3 Dimensionality Mapping

The Euclidean space is assumed to be separated into cells made up of linearly dependent basis called the linear subspace. It is assumed, the projection would map the basis into a linearly independent subspace of lower dimensionality than the linear subspace that spans it. This is analogous to the Generative topographic mapping [4] where the goal is to find a model of the linear subspace. In this new approach, the advantages of an affine transformation would be combined

to that of the projection. The addition of the affine transformation would make the projection degenerate, which would keep the knowledge of the basis secret from the adversary. The rate of approximation of the error vectors from the uniform distribution to random distribution depends on the geometry of the linearly independent subspace. This subspace is embedded tangentially from the linear subspace with complexity of approximation given as $O(c^d)$ where c is the coordinate axis of the linear subspace. For a linear subspace divided into grids of sides L with parameter $2L$ with a constant 2^n, the volume is given by [125]

$$V = 2^n \gamma \left(\frac{n}{2} + 1 \right) \tag{3.10}$$

while the area of the linearly independent subspace is given by

$$\sqrt{(2\pi L)^2 + \gamma^2 (\tilde{B}_i)} \tag{3.11}$$

and the affine transformation is related to the area by

$$f = (Lsin2\pi\tilde{B}, Lsin2\pi\tilde{B}, \gamma B)^T \tag{3.12}$$

where $\gamma \in \{0, \frac{1}{2}\}$. The dimensionality of the space of points that make up the lattice plays a role in the reduction of the basis vectors into its linearly independent variants by deriving a projection matrix that would map the basis vectors accordingly. There is also a corresponding mapping of the diagonal coefficients of the basis vectors to the dimensionality of the reduced linearly independent vectors after it has been projected by the Projection map into a low dimensional grid of lattice points as shown in Figure 3.2.

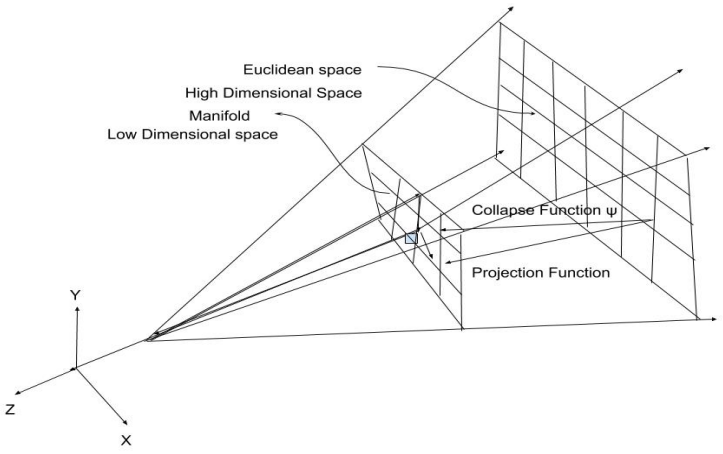

Figure 3.2: Geometric illustration of projection from a high dimensionality plane to a low dimensionality sub-plane

The algebraic structure would be retained to create a tighter reduction to a worst case assumption. In contrast to the LLL reduction approach which reduces the basis to vectors that have an orthogonal property, dimensionality mapping reduces the basis vector to its linearly independent variant with its coefficients being elements of a set of non-negative integers Z at minimal complexity.

There has been several approaches to basis reduction using the dimensionality property. For instance, a randomized variant of Latent factor model was employed to reduce high dimensional sparse matrices[126]. Also, a non linear injective map based on a supervised manifold learning algorithm was constructed to reduce radial basis function[127]. Furthermore, a 2 dimensional unsupervised reduction process was employed to reduce a similarity matrix[128]. Finally, a dimensionality reduction algorithm was proposed to reduce symmetric positive definite matrices belonging to a Riemannian space with variance in the vectors in an unsupervised scenario[129].

Let a basis $B = \sum_{i=1}^{N} b_i$ where $B \in R^n$ connected Z^n space of feasible classes. The goal is to construct a basis in a low dimensional space of points $\tilde{B} = \prod_{i=1}^{N(B_i)} \tilde{B}_i$ which is expressed as linear combination of an affine transformation and a probability distance function. These parameters are expressed as $\tilde{B}_i(x) = P_{ij} f(x)$ where P_{ij} and affine transformation $f(x)$ are defined as $f \colon Z^n \to Z^d$ where Z^d belongs to a low dimensional distribution and $P_{ij} \colon B_i \times \tilde{B}_i \to Z^d$. The probability distance function describes the distance between the basis B_i and its dimensionality reduced variant $\tilde{B}_i$ belonging to different classes. It also computes their dot product respectively. The separation between the classes is expressed using KL divergence[130] $KL(D_q||D_\pi = \int \rho_q log \rho_q / \rho_\pi d\mu$. The affine transformation can also give the Bayesian inference of Z^n. The number of sampled basis $N \subset N(B_i)$ where $N(B_i)$ is the set of the samples belonging to an n-dimensional space with distance from Z^n described by the KL divergence. The affine transformation can be categorized in two types; for points in Z^n and in points Z^{n-m} where Z^m is the space of feasible classes containing $\tilde{B}_i$.

$$fn(B_i, B_j) = \sum_{i=1}^{N} K_i (B_i - B_j)(B_i - B_j) \tag{3.13}$$

$$fn - d(B_i, \tilde{B}_i) = \sum_{i=1}^{N} K_i (B_i - B_j)(B_i - B_j) \tag{3.14}$$

where K_i is the kernel function. The choice of kernel is the sigmoid function of a support vector machine [131] where $K = \tan h(iB_iB_j + \theta)$ which defines the non-

linear separability of the points in Z^m and the points in Z^{n-m}. If this is expanded to P_{ij}, it becomes;

$$P_{ij}(B_i, B_j) = \sum_{i=1}^{N} N_i f_c ||B_i - B_j||_F ||B_i - B_j||_F + \varphi = \tag{3.15}$$

$$\mathtt{Tr}(B^T \Psi B)$$

where φ is a solution to the linear system[132]

$$\left[\begin{array}{c|c} 0 & y^T \\ \hline y & \Psi + \tau^{-1} I \end{array}\right] \left[\begin{array}{c} c \\ \hline \alpha \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline y \end{array}\right] \tag{3.16}$$

$(y_1, \alpha_1), \ldots (y_N, \alpha_N)$ and α_i is a positional scalar.

To generate a maximal variance on the condition of orthonomality and that would achieve uniform ergodicity, (3.13) is expanded as follows

$$f_n(B_i, B_j) = \sum_{i=1}^{N} K_i (B_i - B_j)(B_i - B_j) \tag{3.17}$$

$$= \sum_{i=1}^{N} K_i (\Lambda x_i^{-1} - B_j)(\Lambda x_i^{-1} - B_j)$$

$$= \sum_{i=1}^{N} K_i (\Lambda^2 x_i^{-2} - \Lambda x_i^{-1} B_j - \Lambda x_i^{-1} B_j - B_j^2)$$

$$= \sum_{i=1}^{N} K_i (\Lambda^2 x_i^{-2} - 2\Lambda x_i^{-1} B_j - B_j^2)$$

$$= \sum_{i=1}^{N} K_i (B_i^2 - 2\Lambda x_i^{-1} B_j - B_j^2)$$

$$= \sum_{i=1}^{N} K_i (B_i^2 - 2\Lambda x_i^{-1})$$

$$= \sum_{i=1}^{N} K_i B_i^2 - 2 \sum_{i=1}^{N} K_i \Lambda x_i^{-1}$$

$$= \mathtt{Tr}(B^T B) - \mathtt{Tr}(\Sigma^T \Sigma)$$

Lets define Ψ as a collapse function that transforms and quantifies the endomorphic distance between f_n and f_{n-d} where $\Psi \in Z^{m \times n}$, $B = [\ B_1^T, B_2^T, \ldots B_N^T] \ \in Z^{N \times n}$, given a decay coefficient;

$$\delta = Tr(\partial \Sigma^T D^*) + ||P_{ij} - x_i|| + ||P_{ij} - \tag{3.18}$$

$$x_j|| + ||B_i||$$

further expansion of the collapse function becomes

$$\Psi = \min_{f_n,i,j} \sum_{i=1}^{N} f_n(B_i, B_j)(\mathrm{Tr}(\partial H^T D^*) + ||P_{ij} - x_i|| + ||P_{ij} - x_j|| + ||B_i|| \quad \text{subject to} B^T P_i B = I, \partial \Sigma^T P_j \partial \Sigma = I \tag{3.19}$$

where P_{ij} is a sparse approximation of a projection function P and $\tilde{B} = fP$ which implies that $\mathrm{Tr}(\tilde{B}^T P_{ij} \tilde{B}) + ||f||_F$. This further confirms the fact that the variance of the Hessian matrix as regards to the Newton Direction, defined by the KL divergence must meet the conditions of orthonomality of the subset sum of the basis and the dimensionality of the projection mapping which will in turn lead to resulting low dimensionality vectors being linearly independent. The objective of the reduction process is to minimize the class factor of the collapse function Ψ. The optimization problem is formulated as

$$\min_{\Psi} \max_{f} \mathrm{Tr}(\tilde{B}^T \Psi \tilde{B}) + ||f||_F$$

$$\text{subject to} B^T P_i B = I \& \partial \Sigma^T P_j \partial \Sigma = I$$

An arbitrary noise matrix $W \in R^n$ is chosen. This noise term is equivalent to the dimensionality of the space that contains the dimensionality reduced basis. This would expand the Newton Direction function, which is a function of the log likelihood ratio. This ratio realigns the direction of the collapse function from the linear subspace to the low dimensional linearly independent subspace, the noise term and is given as

$$D^* = \arg\min \log \int \frac{\rho_q}{\rho_\pi} d\mu + \mathrm{Tr}(W^T B_i W) + \frac{1}{2} \sum_j f b W^T \tag{3.20}$$

It can be seen that realignment is enhanced by the orthonormal columns of the arbitrary noise matrix W. To further test for invertibility, a regularization parameter is employed as follows

$$\max\mathrm{Tr}[\ (f_n + DI_\mu)^{-1} f_{n-d}]\ \leq \Delta \tag{3.21}$$

where Δ is a regularization parameter and I_μ is the rate function . The overall

optimization problem becomes

$$\min_{\Psi}\max_{f}\sum_{i=1}^{N(x_i)}\mathrm{Tr}(B^T\Psi B)f_n(B_i,B_j)(\mathrm{Tr}(\partial H^T D^*)+$$

$$||P_{ij}-x_i||+||P_{ij}-x_j||+||B_i||)+\mathrm{Tr}(\tilde{B}^T\Psi\tilde{B})+||f||_F$$

$$\text{subject to} B^T P_i B = I,$$

$$\partial\Sigma^T P_j \partial\Sigma = I$$

$||P_{ij}-x||$ can be solved using lemma 3[133]

$$\sum^{N(x_i)}||P_{ij}-x_i|| = min||P-B||_F^2 \tag{3.22}$$

$$= min||P^2||_F + \min_{x\in Z/\Lambda} 2x^T P + ||B||_F^2$$

then the trace samples of the input point is solved as follows

$$\mathrm{Tr}(B^T\Psi B)=\sum_{i,j}^{N} f(B_i,B_j)+\Psi(log(B_i)-(log(B_j)) \tag{3.23}$$

Finally, the trace samples of the product of the hessian matrix and the Newton Direction is solved using lemma 2[134]

$$minmin\mathrm{Tr}(\bigtriangledown H^T D^*) \leq -vec(D^{*T}B^T sign(\sigma)vec(D^*)- \tag{3.24}$$

$$||P_{ij}+D^*||_F+||P_{ij}||_F$$

$$\text{subject to} D^* P D^* = I$$

$$\text{where } sign(\sigma)\begin{cases} =1 & \text{if } r_i>0 \\ \in[-1,1] & \text{if } r_i=0 \\ =-1 & \text{if } r_i<0 \end{cases}$$

Algorithm 6 Dimensionality Mapping

Require: Basis $B=(b_1,\dots b_n)\in Z^{n\times n}$, Random Matrix $W\in R^n$,Affine Transformation f

Ensure: Dimensionally Reduced Basis $\tilde{B}=\tilde{B}=\{\tilde{b}_1,\dots,\tilde{b}_n\}\in Z^{n\times m}$

1: $D=Tr(W^TBW)+\frac{1}{2}\sum_j fbW^T$
2: **for** $i,j\leftarrow n$ **do**
3: $ComputeEmbeddingP_{ij}(X_i,X_j) = \sum N_i f||X_i-X_j||_F.||X_i-X_j||_F+ = \mathtt{Tr}(B^T\Psi B)$
4: $SolveOptimization\ Tr(\tilde{B}^T\Psi\tilde{B})+||f||_F$
5: $checkforConvergence$
6: Return $\tilde{B}$

In line 1, computing the trace of the weight matrix would require n^2 scalar operations and adding the noise term would amount to $O(n^2log_2 n)$ operations. Computing a sparse approximation of the projection as a function of the affine function

would result to $O(n^2)$ operations while in line 5, computing an approximation of the optimization problem would result to $O(log_2 \frac{B}{n})$ operations. This gives a total complexity of $O(n^2 + n^2 + log \frac{B}{n})$

3.3.1 Ergodicity

In this section, the extent to which the lattice reduction process from dimensionality mapping achieves uniform ergodicity to the theoretical approximation of the shortest vector problem is analyzed. The Lebesgue measure is chosen as $\mu \in (0, \frac{1}{2})$. It is expected that after an infinite number of iterations, the asymptotic distance between the high dimensional space of the distribution converges on the condition that the concatenation of the projection matrix and the affine transformation generate a product that would force the collapse function to return a solution. This solution maps the vectors from linear dependency to linear independency. In other words, satisfy this condition $\lim_{i \to \infty} ||\lambda_{k,i} - x_{k,i}|| = 0 \forall k, i \in F_q$, if $\lim_{i \to \infty} ||\lambda_i - P_{ij}(f_j)|| = 0$. In clear terms, uniform ergodicity can be achieved if the vector are linearly independent. It can be shown that the L_2 norm of the error between the sparse approximation of the projection operator and the lattice basis can be defined in terms of the affine transformation f and the preimage of the lattice basis $x_{k,i}$. Inspired by the approach [135], this becomes,

$$||Pij - x_i|| = ||B_i(x_{k,i} - \mu_{k,j} \frac{f_{k,i}(x_{k,i})}{||f'_{k,i}(x_{k,i})||^2} f'_{k,i}(x_{k,i})) - x_{k,i}|| \tag{3.25}$$

$$= ||B_i x_{k,i} - \mu_{k,j} \frac{f_{k,i}(x_{k,i}) B_i}{||f'_{k,i}(x_{k,i})||^2} f'_{k,i}(x_{k,i})) - x_{k,i}|| \tag{3.26}$$

$$\leq ||\Lambda_{k,i} - \mu_{k,j} \frac{f_{k,i}(x_{k,i})}{||f'_{k,i}(x_{k,i})||^2} f'_{k,i}(\Lambda_{k,i}) - x_{k,i}|| \tag{3.27}$$

$$= ||\mu k, j \frac{f_{k,i}(x_{k,i})}{||f'_{k,i}(x_{k,i})||^2} (\Lambda_{k,i})|| \tag{3.28}$$

$$= \frac{||\mu_{k,j} f_{k,i}(x_{k,i})(\Lambda_{k,i})||}{||f'_{k,i}(x_{k,i})||} \tag{3.29}$$

$$= \mu_{k,j} \frac{\langle f_{k,i}, \Lambda_{k,i} \rangle}{||f'_{k,i}||_F} \tag{3.30}$$

This shows that the error and linearly independent vectors results in an affine transformation that is orthogonal. This affine transformation can be said to be an image of the original lattice basis. If the condition of (3.15), $B^T P_i B = I$ holds true that means that $\lim_{i \to \infty} B^T P_i B - I = 0$. This further becomes $\lim_{i \to \infty} ||\Lambda_{k,i} - x_{k,i}|| = 0$. Therefore,it follows that if $||\Lambda_{k,i} - x_{k,i}|| = B^T P_i B = I$, the expected solution would

converge at some point. If an update parameter is constructed to force the collapse function to return a solution as the iterations increases in size, it can be seen that the Newton direction moves in tandem with the regularization parameter in (17),

$$||\Lambda_{k,i} - P_{ij}\bigg(log \int \frac{\rho_q}{\rho_\pi} d\mu + \mathrm{Tr}(W^T B_i W) + \frac{1}{2}\sum_j fbW^T\bigg)|| \tag{3.31}$$

$$= ||\Lambda_{k,i} - P_{ij}\bigg(log \int \frac{d\mu}{\rho_\pi} \rho_q + B_i(W \oplus W) + \frac{1}{2}\sum_j fbW^T\bigg)|| \tag{3.32}$$

$$\leq ||\Lambda_{k,i} - E_\mu[\ \langle \omega_\mu(x_{k,i}), v(x_{k,i})\rangle]\ log\frac{\rho_q}{\rho_\pi} + (\frac{1}{2}\sum_j fb + B_i)W^2)|| = ||\Lambda_{k,i} - E[\ ||\omega_\mu(x_{k,i}) - v(x_{k,i})||^2]\ log\frac{\rho_q}{\rho_\pi} + \frac{W^2}{2}(\sum_j fb + B_i)|| \tag{3.33}$$

$$= ||\Lambda_{k,i} + \frac{\partial^2 J(x_{k,i})}{\partial^2 H_{\omega,k}} \frac{\rho_q}{\rho_\pi} + \frac{W^2}{2}(\frac{1}{2}\sum_j fb + B_i)|| \tag{3.34}$$

where $J(x_{k,i})$ is a special convex set that defines the convex property of the Projection map and can be expressed in terms of the update term as follows

$$J(x_{k,i}) = \frac{\upsilon_{k,i}^T}{\Delta}(\sum_{j=1}^{n} \alpha_j \Psi_j) - (\frac{1}{n}\sum_{i=1}^{n} \Psi_i) \tag{3.35}$$

where α_j is a smoothing parameter and $\upsilon_{k,i}^T$ determines the geometry of the linear subspace. The role of the smoothing parameter is to limit the amount rate of change of the Legendre measure which would affect the divergence between classes and in the process keep the dimensionality reduced basis in a fixed point. The probability density function of the theoretical distribution can be expressed as a function of the absolute value of a covariance matrix Σ which is usually expressed in relation to a transition matrix T whose column is orthonormal as follows

$$log\rho_q = -\frac{1}{2}log((2\pi)^d|\Sigma|) - \frac{1}{2}(x-\mu)^T(\Sigma)^{-1}(x-\mu) \tag{3.36}$$

3.3.2 Experiment

The ideal lattice chosen are usually subsets of Z^n. At high dimensionality, samples of the subsets is projected to exponentially many lattice vectors which leads

to losses in the signal-to-noise ratio. The total number of vectors that leads to this losses is given by [136]

$$N_k = \sum_{k=1}^{k} SNR^{\frac{\epsilon T}{k}} \tag{3.37}$$

where $\frac{\epsilon T}{k}$ is the multiplying gain of the MIMO detector. The logarithm of the orthogonal diagonal coefficient as a function of Equation 3.10 can be used to define and parameterize the quality of basis [118]. To analyse the quality of the reduced basis using dimensionality mapping method, its Bit Error Rate(BER) performance when applied to reduce a parity check matrix is presented. This lattice reduction approach is applied to a MIMO(Massive Input Massive Output) detector systems. Its BER performance is compared with other lattice basis reduction methods. The signal-to-noise ratio is given as $S_{NR} = N_T * E_S/N_o$ where the average bit is E_S. A good performance implies the capability to approximate the basis vectors to a uniform distribution in essence solving the shortest vector problem with lesser complexity. The MIMO detector systems used are the 4×4, 6×6, 8×8 uncoded system with 4QAM, 16QAM and 64QAM constellation in a Zero-Forcing(ZF) detector system. The lattice reduction method considered are the LLL, Modified-LLL(MLLL)[137], Gram Schmidt Orthogonalization based reduction ,QR decomposition. During the process of reducing the basis, the symbols of the constellation are shifted and scaled by an integer orthogonal matrix before projecting the zero forcing vector back to the constellation [138]. For a given signal-to-noise ratio, the projections are sparse and random at low dimensionality. For both the LLL and MLLL method, the trade-off factor δ is assigned a value $\frac{3}{4}$. The lattice based channel matrix used as basis in the simulation had entries from a complex i.i.d discrete Gaussian distribution with regularization parameter $\mu = 0$, variance $\sigma = 1$ generated over 100000 Monte Carlo runs. In addition, for the QR decomposition reordering the columns of the generator matrix of the lattice downgrades its performance overtime. From the Results, it can be seen that the Modified-LLL(MLLL) gives a better BER performance than dimensionality mapping due to the number of column swaps in the MLLL method. Furthermore, the condition of using the orthogonality of the basis vectors with the Projection as a means of finding a solution to the optimization problem impacts on the the BER performance of using dimensionality Mapping. However, dimensionality reduction gives a performance improvement of about 1db on the 4×4 uncoded system using 4QAM constellation over the LLL algorithm for a Bit error ratio of 10^{-4} as shown in Figure 3.4 and a performance improvement of about 4db on the

6×6 uncoded system using 4QAM constellation over the LLL algorithm for bit error ratio of 10^{-4} as shown in Figure 3.5. It also outperformed other basis reduction methods; Orthogonalization and QR decomposition. It can be said that at lower signal-to-noise ratio, the optimization problem that is formulated to extract the mapping function is prone to more unknowns. This would lead to complexity in convergence thereby making the proposed method not ideal for reducing basis. It can be inferred that the proposed dimensionality mapping/reduction supports linear transformation rather than unimodular transformation. Also, the bit error rate result can be attributed to the proposition that the dimensionality mapping processes a certain number of independent vectors of the lattice that varies in a given permutation without breaking the constraints of the optimization problem. Furthermore, it would be stated that LLL performance degrades with increase in the orthogonality factor at high dimensional subspaces [139]

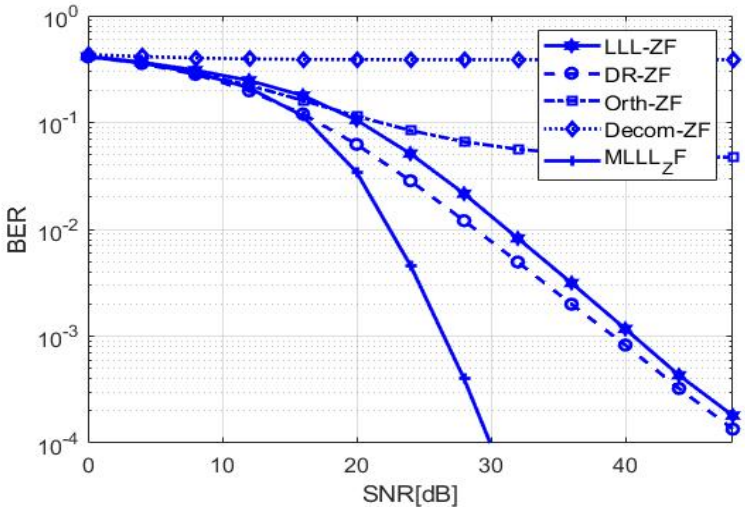

Figure 3.3: Bit Error Rate of various Basis reduction methods for Gaussian Sampler applied to ZF-MIMO detectors in 4×4 uncoded system using 16QAM constellation

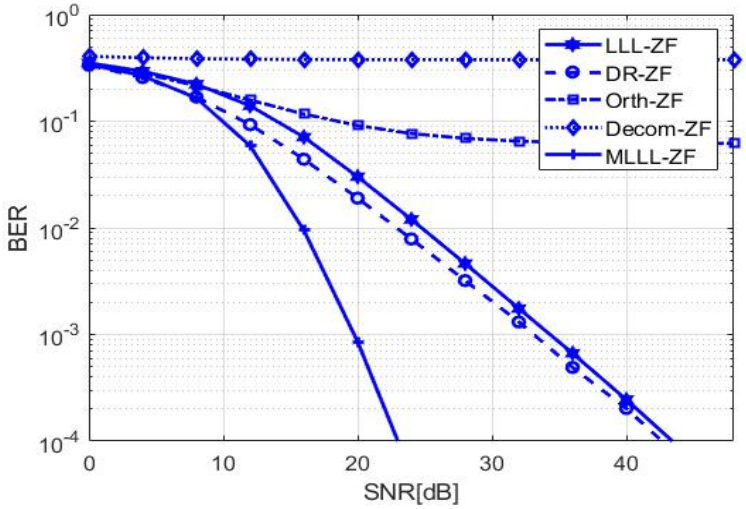

Figure 3.4: Bit Error Rate of various Basis reduction methods for Gaussian Sampler applied to ZF-MIMO detectors in 4×4 uncoded system using 4QAM constellation

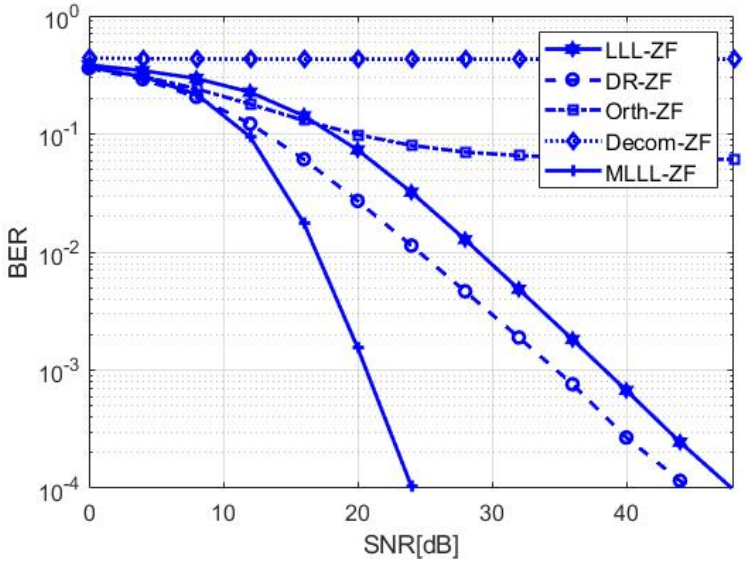

Figure 3.5: Bit Error Rate of various Basis reduction methods for Gaussian Sampler applied to ZF-MIMO detectors in 6×6 uncoded system using 4QAM constellation

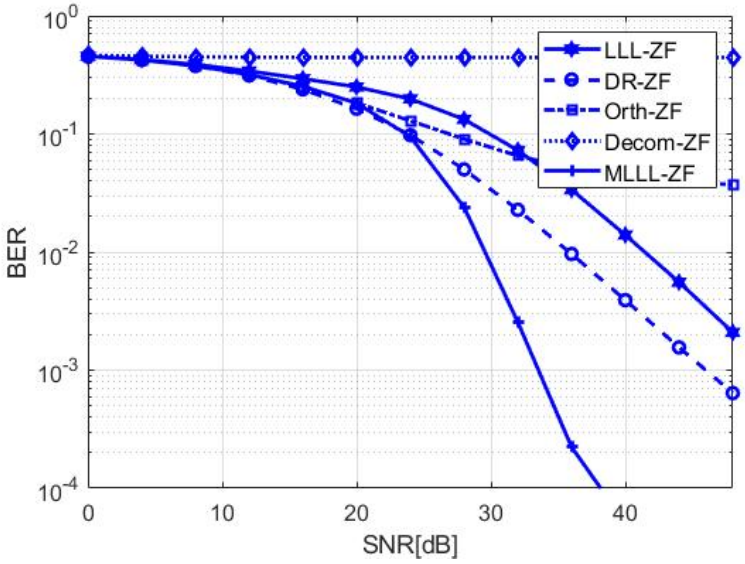

Figure 3.6: Bit Error Rate of various Basis reduction methods for Gaussian Sampler applied to ZF-MIMO detectors in 6×6 uncoded system using 64QAM constellation

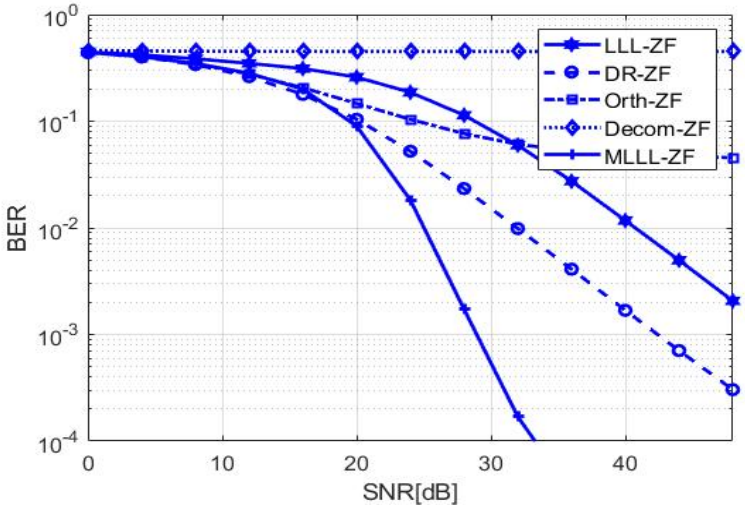

Figure 3.7: Bit Error Rate of various Basis reduction methods for Gaussian Sampler applied to ZF-MIMO detectors in 8×8 uncoded system using 16QAM constellation

3.4 Gaussian Sampler

A modified variant of the nearest plane approach is given in Algorithm 7. The single value decomposition is employed to reduce the covariance matrix Σ expressed as $\Sigma_1 = \sigma^2 DCD^T$ where D is a diagonal matrix and C a correlation

coefficient matrix. The lattice points are sampled from an ellipsoidal discrete Gaussian distribution with covariance $\sigma^2 DD^T C$. This makes the resulting sample z to be sampled in a spherical discrete Gaussian distribution, thus making the complementary covariance matrix to become $\Sigma_2 = \sigma^2 I$. Then it is translated by adding a small distribution with covariance $(\sigma^2 I - \sigma_2 DCD^T) + \sigma^2 DD^T C = \sigma^2$ for some $\sigma = \omega(\sqrt{\log n})$. In [26] the bound on the diagonal matrix D was stated as $(2(D)+1)$, which results to the positive definite covariance matrix $\Sigma_1 = \sigma^2 DCDT = \sigma^2$ and $\Sigma_2 = \sigma^2 I + \sigma^2 DCD^T = \sigma^2(DCD^T + I)$. The matrix D and C should have small entries to enable easier computation and eliminate the need for offline computation. The first $k-1$ columns of C are orthogonal basis of Σ_1 column space. Consequently, the norm of Σ_1 approaches σ^2 as k tends to infinity. The orthogonal structure of C is in far contrast to the triangular structure because reduction with a triangular structure is not straightforward making computation exponentially high on input.

Algorithm 7 Gaussian Sampler

Require: Basis $B = (b_1, \dots b_n) \in Z^{n \times n}$, parameter σ, target center $c \in Z^n$, Low dimensional basis $\tilde{B} = \{\tilde{b}_1, \dots, \tilde{b}_n\} \in Z^{n \times m}$, norm $r_i = \|b_i^*\|$

Ensure: z_i drawn from a distribution statistically close to $D_{\Lambda,\sigma,c}$

$v_n \leftarrow (0, \dots, 0), c_n \leftarrow c$

for i $\leftarrow n, \dots, 1$ **do** $\quad t_i = \frac{\langle c_i, \tilde{b}_i \rangle}{\|b_i\|^2} - \sum_{j>i}^{m} a_{i,j} z_j - \sum_{j'>i}^{n} a_{i,j'} z_j$ where $a_{i,j}$ is an integer coefficient close to c_i

    compute a covariance and complementary covariance matrix as a function of Gaussian parameter $\Sigma_1 = \sigma^2 BBC^T, \Sigma_2 = \sigma^2 I + \Sigma_1 = \sigma(BCB^T + I)$

    $z_i \leftarrow D_{Z,\sqrt{\Sigma_i},t_i}$

    $t_{i-1} \colon (\alpha_{i,j} + z_i) b_i^* \leftarrow t_i$

    $t_{i-1} \colon (\alpha_{i,j} + z_i) b_i^* \leftarrow Z^n$

    $v_{i-1} \colon z_i b_l^* \leftarrow v_i$

    Calculate probability rate

    if $r <$ probability rate **then**

        Set $y = t_0 - v_0$

        Return y

It is very important that the statistical distance between the proposed distribution where the lattice points are sampled from and the target distribution where the lattice points are ideal be minimized as possible. This can also be described using the concept of convergence. When the range of the center c of the distribution is denoted by $[\ -0.5, 0.5]$ and $\sigma > 0$, the probability rate $\lambda(x, y)$ increases for c fixed at $[\ 0, 0.5]$ [140]. In this proposed design $c \in Z^n$ would be chosen at $[\ 0, 1]$. The input to the proposed sampler would be a basis B, Gaussian parameter σ, the center

of the distribution $c \in Z^m$, the initial sample x of the Markov chain and the tail cut parameter τ. A new sample would be generated if U a random number generated from a uniform distribution $[\ 0, 1)$ meets the following condition $U > \oslash[\ t]$ where $\oslash[\ t]$ is the cumulative distribution function and t is the result of the reduced basis $\tilde{b}_i$ and the center with a norm $\|\tilde{b}_i\|$. The final sample is accepted if $U \leq \lambda(x, y)$. Based on the Gaussian sampler, the hardness of the learning with error problem can be redefined to finding $v \in \Lambda$ such that the distance between c and $(z_i b_i^*)$ is short and also $|c_i - t_i| \leq |t_0 - v_0|$. This is because the target vector is sampled from the required distribution $z_i \leftarrow D_{Z,\sqrt{\Sigma_i},t_i}$ they are spaced at the norm r_i. Also, $|c_i - t_i|$ is the distance of c_i from a plane in the distribution. From Babai's theorem, $|c_i - t_i| \leq \frac{\|b_i\|^2}{2}$. If $z \leftarrow D_{Z,\sqrt{\Sigma_i},t_i}$, then $c_0 - v_0$ is close to $z_i b_i^*$. Therefore,

$$|c_i - t_i| = |c_i - y + v_0| \leq \xi(t_i - z_i) \leq \xi|c_i - z_i|\frac{\|b_i\|^2}{2} \tag{3.38}$$

$$\begin{aligned} |c_i - t_i| &= (|t_i - t_{i-1}|^2 + |c_i - t_i|^2 \leq |c_i - z_i|(1 + \xi) \\ &< \xi(c_i - z_i) \end{aligned} \tag{3.39}$$

Lemma 9. *Let B be a basis of a lattice, let $\Sigma = DD^T\sigma^2$ for $\sigma > 0$ and $c \in Z^n$ over dimensionality Reduced basis $\tilde{B}$ to generate convolute t. The probability of outputting v has distribution statistically close to $D_{Z,\sqrt{\Sigma},t_i}$.*

Proof. Let $\Sigma = BB^{-1}\sigma^2$ for $\sigma > 0$, the probability that $z_i = \bar{z}_i$ is satisfied if $\Lambda + t_i = \bar{x}$ where $\Lambda + t_i$ is the support of $y = x - v_0$ which is a solution to the shortest vector problem. This resolves to $y = \bar{x} - z_i bi$. From Lemma 3[26], $\bar{x} = t - B\bar{z}$, which results to $y = t - B\bar{z} - \bar{z}BB^{-1}$. This satisfies the Gaussian distribution given as $\rho_{\sqrt{\Sigma_i},t_i}(\bar{x} - z_i b_i) = \rho_{\sqrt{\Sigma_i},t_i}(t - B\bar{z} - \bar{z}BB^{-1}) = \rho_{\sqrt{\Sigma_i},t_i}(B\bar{z}) = \rho_{\sqrt{\Sigma_i},t_i,Z}(v)$

□

3.4.1 Statistical distance

This algorithm as expected samples the entire variable as a block z_i. This is because sampling a component at each step leads to inefficiency towards closing the gap in the statistical distance. Inspired by the convolution theorem [26], Theorem 3 is used to test for correctness of the constructed sampler. Readers are encouraged to reference it for more details

Theorem 3. *For $\epsilon \in (0, \frac{1}{2})$, let Σ_1 and Σ_2 be a positive definite matrices where $\Sigma_1 = \sigma^2 BB^T C$ and $\Sigma_2 = \sigma^2(BCB^T + I) \in \eta_\epsilon^2(\Lambda_1 + \Lambda_2)$, $\Lambda_1(B) \in Z^n$ and $\Lambda_2(B) \in Z^m$.*

Given a basis $B = \{b_1, \dots b_n\}$ *and with two arbitrary distributions* $x_2 \leftarrow D_{\Lambda_2, c_2, \sqrt{\Sigma_2}}$ *and* $x_1 \leftarrow D_{\Lambda_1, \sqrt{\Sigma_1}}(c_1 + x_2)$ *that produces a vector* $x = (x_1, x_2) \in (\Lambda_1 + \Lambda_2)$,*the statistical distance* $D_{\Lambda_2 + c_1, x}$ *is bounded by* $1 + 4\epsilon$.

Proof. It would be shown that the output distribution of the sampler is statistical close to the target discrete Gaussian distribution, $x_2 \leftarrow D_{\Lambda_2, c_2, \sqrt{\Sigma_2}}$ and $x_1 \leftarrow D_{\Lambda_1, \sqrt{\Sigma_1}}(c_1 + x_2)$ which follows that $Pr[\ x_1 = x_{/,1}, x_2 = x_{/,2}]$

$$\frac{\frac{\frac{\rho_{\sqrt{\Sigma_1}}(x_1 - c_1).\rho_{\sqrt{\Sigma_2}}(x_2 - c_2)}{\rho_{\sqrt{\Sigma_1}}(\Lambda_1 + BC^{-1}(c_1 + x_2).\rho_{\sqrt{\Lambda_2}(x_2 - c_2)}}}{\frac{\rho_{\sqrt{\Sigma_1}}(x_1 - c_1).\rho_{\sqrt{\Sigma_2}}(x_2 - c_2)}{\rho_{\sqrt{\Sigma_1}}(\Lambda_1 - c_1 - BC(c_1 - x_2).\rho_{\sqrt{\Lambda_2}(x_2 - c_2)}}}}{\frac{\rho_{\sqrt{\Sigma_1}}(x_1 - c_1).\rho_{\sqrt{\Sigma_2}}(x_2 - c_2)}{\sum_{x_2 \in \Lambda_2} \rho_{\sqrt{\Sigma_1}}(x_2 - c_1) + BC(c_1 - x_2).\rho_{\sqrt{x_2}(x_2 - c_2)}}}$$

$$\frac{\frac{\rho_{\sqrt{\Sigma_1}}(x_1 - c_1).\sum_{x_2 \in \Lambda_2}(\Lambda_2 - c_2)}{\sum_{x_2 \in \Lambda_2} \rho_{\sqrt{\Sigma_1}}(\Lambda_2) + BC(c_1 + \Lambda_2).\rho_{\sqrt{x_2}(x_2 - c_2)}}}{\propto \frac{\rho_{\sqrt{\Sigma_1}}(x - c_1).[\ 1, \frac{1-\epsilon}{1+\epsilon}]}{\sum_{x_2 \in \Lambda_2} \rho_{\sqrt{\Sigma_1}}(\Lambda_2) + BC(c_1 + \Lambda_2).\rho_{\sqrt{x_2}(x_2 - c_2)}}}$$

$$\in \left[\ \frac{1-\epsilon}{1+\epsilon}\right], \left[\ \frac{1-\epsilon}{1+\epsilon}\right] . \frac{\rho_{\sqrt{\Sigma_1}}(x - c_1)}{\rho_{\sqrt{\Sigma_2}}(x - \Lambda_2)}$$

$$\left[\ \frac{1-\epsilon}{1+\epsilon}\right] \left[\ \frac{1-\epsilon}{1+\epsilon}\right] . D_{\Lambda_2, c_1}(x)$$

$$[\ 1 - 4\epsilon, 1 + 4\epsilon]\ . D_{\Lambda_2, c_1}(x)$$

□

The preimage of the target center c would be computed after z is sampled from a distribution statistically close to $D_{Z, \sqrt{\Lambda_i}, t_i}$. The required sample y from the sampler shown is accepted if the probability of acceptance $\lambda(x, y)$ is minimal.

Theorem 4. *Let* $\Lambda(B) \in Z$ *and* $\epsilon \in (0, \frac{1}{2}]$. *Let there be a support* $\Lambda + c$ *where* $|x - c| < k\sigma$ *for* $k > 0$, *the bound on the probability rate of accepting a sample* $y = x - v_0$ *from the discrete Gaussian sample is given as* $min\left\{1, \frac{D_{\Lambda + c, \sigma(y)}}{D_{\Lambda + c, \sigma(x)}} . \frac{D_{Z, \sigma_i, y(x)}}{D_{Z, \sigma_i, x(y)}}\right\} \leq 16\epsilon + 8e^{-\frac{k^2}{2}}$

Proof. The probability of acceptance is given by $min\left\{1, \frac{D_{\Lambda + c, \sigma(y)}}{D_{\Lambda + c, \sigma(x)}} . \frac{D_{Z, \sigma_i, y(x)}}{D_{Z, \sigma_i, x(y)}}\right\}$ which follows that

$$min\left\{1, \frac{D_{\Lambda+c,\sigma(y)}}{D_{\Lambda+c,\sigma(x)}}.\frac{D_{Z-y,\sqrt{\Sigma_i},y-c}}{D_{Z-x,\sqrt{\sigma_i},x-c}}\right\}$$

Expanding the distribution into its constituent Gaussian functions

$$= min\left\{1, \frac{\rho_{\sigma,0}(\Lambda)}{\rho_{\sigma,c}(y)}.\frac{\rho_{\sigma,c}(x)}{\rho_{\sigma,0}(Z)}.\frac{\rho_{\sqrt{\Lambda},c}(y-x)}{\rho_{\sqrt{\Lambda},t_i}(\Lambda)}.\frac{\rho_{\sqrt{\Sigma},t_i}(Z)}{\rho_{\sqrt{\Sigma},c}(x-y)}\right\}$$

Rearranging likely terms

$$= min\left\{1, \frac{\rho_{\sigma,0}(\Lambda)}{\rho_{\sqrt{\Sigma_i},t_i}(\Lambda)}.\frac{\rho_{\sqrt{\Sigma_i},t_i}(\Lambda)(Z)}{\rho_{\sigma,0}(Z)}.\frac{\rho_{\sigma,c}(x)}{\rho_{\sigma,c}(y)}.\frac{\rho_{\sqrt{\Sigma},c}(y-x)}{\rho_{\sqrt{\Sigma},c}(x-y)}\right\}$$

which satisfies

$$min\left\{1, (1+4\epsilon+8e^{-\frac{k^2}{2}})^2(1+4\epsilon^2)\right\}$$

since $k \leq \infty$ and ϵ is $ngl(n)$ then

$$min\{1, (1+16\epsilon+8e^{-\frac{k^2}{2}})^2(1+4\epsilon^2)$$

□

3.4.2 Precision analysis

In this section, the bound on the floating point precision is shown, if the adversary makes q private key queries to the challenger with a reduction factor ϵ. This reduction factor enables the discrete probability density functions to be stored in a fixed table size in memory. This reduction factor is lower bounded as follows $|\epsilon_i| \leq 2^{-m}\rho_{\sigma,0}$ [141]. Employing [111] theorem ,the correctness of precision can be tested as follows

$$k = m^{'} - (C\epsilon + 2log_2(\|B^{-1}\|) + log_2(\mu^2 n^3 \tau\sigma^2 q)$$

$$k = m^{'} - \left(\frac{\sqrt{log(2k(1+\frac{1}{\epsilon})}}{\pi}\right) + 2log_2(\|B^{-1}\|) + log_2(\mu^2 n^3 \tau\sigma^2 q)$$

$$k = m^{'} - \left(\frac{\sqrt{log(2k(1+\frac{1}{\epsilon})}}{\pi}\right) + 2log_2\sqrt{n}.$$

$$\sqrt{\sum_{i\in[\ n]} g_{i,i}^2} + log_2(\mu^2 n^3 \tau\sigma^2 q)$$

$$= m^{'} - \left(\frac{\sqrt{log(2k(1+\frac{1}{\epsilon})}}{\pi}\right) + 3log_2.\sqrt{\sum_{i\in[\ n]} g_{i,i}^2}(\mu^2 n^{3/2} \tau\sigma^2 q)$$

$$k = m^{'} + 3log_2\sqrt{n}.\sqrt{\sum_{i\in[\ n]} g_{i,i}^2} + 3log_2(\mu^2 n^{3/2} \tau\sigma^2 q)$$

To ensure λ-bits of security, the floating point precision should be bounded by $m^{'} \geq \lambda + log_2 \sum g_{i,i}(\mu^2 \tau sqn^{3/2})$ where μ^2 is the precision of the perfect sampler and n is the number of samples from the uniform vectors. Table 3.1 compares the bound on precision.

Table 3.1: Bounds on the Precision of Floating point arithmetic

Bound	
$m \geq \log(lmq/\epsilon)$	[141]
$m \geq \log(lmqe^{-2\pi/s^2}/\epsilon) + 1$	[141]
$m \geq \epsilon + 2.3 + log_2(smq)$	[103]
$m \geq \epsilon + log_2 \sum g_{i,i}(\mu \tau sqn^{3/2})$	This work

3.4.2.1 Simulation result

The performance of the precision of the sampler was tested with 10^5 samples. These samples is assumed to guarantee the success probability. The plot is shown in Fig 3.8. The parameters employed where from the BLISS[142] signature scheme. The tail cut parameter is expressed as $\tau = 13.4/\sqrt{2\pi} = 5.36$, the precision of the perfect sampler was taken as 1024 bits, the Gaussian parameter was taken as 254. Consequently, computing the standard deviation becomes $s = \sqrt{2\pi}215.75 = 541$. Also, the success of probability of 2^{-128} for a security level of 2^{256} was taken into consideration. Finally, adversary adaptive queries to an oracle was given as $q = 2^{64}$.

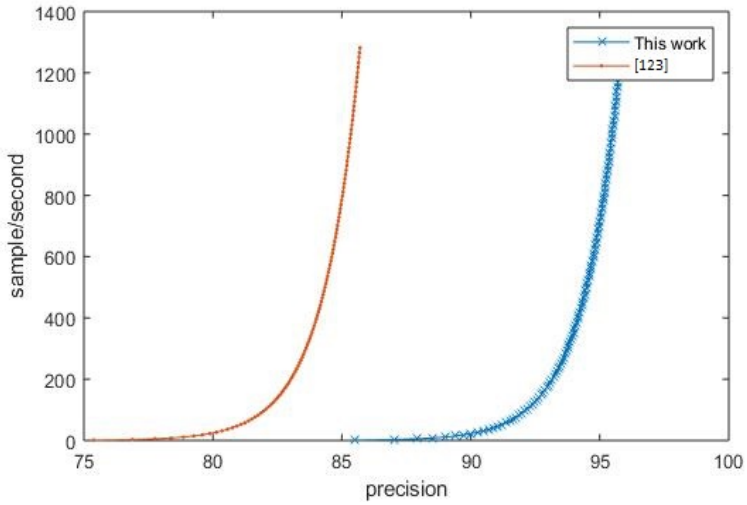

Figure 3.8: Precision of Gaussian sampler for 1400 samples per second, tail cut parameter $\tau = 5.36$, $\sigma = 541$, adaptive queries $q = 2^{64}$

From the result it can be observed that at a sampling size of 10^{20}, a precision of 105 bits was recorded. For the same sampling size, [141] recorded a precision of 188 and 180 bits while [103] recorded a precision of 100 bits. The precision is inversely proportional to the success probability of the system withstanding an attack and directly proportional to the security parameter.

Chapter Summary

Gaussian sampling is one of the schemes that are employed in Lattice cryptography to approximate a vector with the shortest norm in Euclidean space. In other words to solve the shortest vector problem in a lattice. This method is a basis for the Learning with errors concept which is intractable in polynomial time. In order to approximate a solution to this problem, a Gaussian sampling pre-processing method where a basis is mapped from a high dimensional subspace to a Low dimensional subspace using a projection function is proposed. This method is analogous to the LLL method which reduces the basis to its Gram-Schmidt orthogonal equivalent with its attendant complexity. The security of these scheme is based on the adversary's lack of knowledge of the basis that generated the vectors. It is assumed the Gaussian distribution has the probability mass centred around vectors with huge norms. The random sampling of the basis vectors from a distribution depends on the geometry of linear subspace. The algebraic structure of the lattice points is still maintained for reduction to worst case assumptions using random oracles. The sampled vectors is a subset of the n-dimensional space. Its statistical distance is parameterized by the KL divergence.In order to further this approach, a "collapse" function was constructed as an optimization problem together with an affine transformation that would enable a smooth projection. BER performance of the reduced basis show that the proposed method outperforms the LLL method. However, with slightly less than the expected result compared the modified variant. Pierkert's Convolution theorem was extended to improve on Poppleman's approach to investigate the correctness of a proposed variant of Babai's plane algorithm where the next iteration of the center is a mapping process. On input of success probability 2ϵ, the statistical distance between the simulated distribution and an ideal distribution is reduced from 32ϵ to 2ϵ. Also tighter bounds on the precision of the distribution was derived and simulated using parameters from BLISS signature [142]. This shows our approach is efficient for algorithmic complexity purposes.

Chapter 4

Syndrome Decoding problem: Solution using Plücker coordinate of the Grassmannian

4.1 Introduction

In the key generation process of code based cryptosystem, the generator matrix G is scrambled to $G^{'}$ using a combination of permutation matrices. Given a message bit to be encrypted m, the generated ciphertext can be represented as a syndrome $s = mG^{'} + x$. The vector x indexes the columns of the parity check matrix H of the code. In other words, $Hs^T = H(mG^{'} + x) = Hx^T$. The goal of the adversary is to recover the vector x of minimum weight w, then G can be extracted using some form of interpolation thereby revealing m. To compute the vector x given the syndrome s and the parity matrix H is an NP complete problem termed the syndrome decoding problem. Most approaches to solve this problem is to first lift this vector into a basis of a vector space. Then, construct a set of indices containing linearly independent columns of the parity matrix [143] or create a support containing the minimum codewords of a certain rank [10]. The dimension of the set corresponds with the dimension of the code. To prune the set, Gaussian elimination is employed. For the Grassmann, the parity matrix are parameterized by the Plücker coordinates. In addition, the scrambled generator matrix is a matrix representation of an isomorphism. In other words, the isomorphism hides the structure of the message symbol and in the process forms a code with minimum distance d_{min}. The minimum Grassmann weight of the vector x is given as $G_{wt}(x) \leq \frac{d_{\min}}{2}$.

Furthermore, the vector would be lifted as a basis of a subspace of a projective space. Therefore, the solution to the syndrome decoding problem reduces to a situation of creating set of linearly independent coordinates and iterating over this set to find a minimum distance $d_{min} \leq n - k + 1$ of a lifted code. This thesis is more especially interested in the probabilities of enumerating this basis to find linearly independent Plükcer coordinates that will compute this minimum distance .

The hardness of decoding the syndrome of a linear code [144] has been useful in designing quantum safe encryption in the Hamming metric using Goppa codes [145] and in the rank metric using Gabidulin codes [67]. The syndrome decoding problem is the basis of cryptanalysis in code based cryptography. This is because on the input of certain code parameters and with the knowledge of the structure of the code, an attacker can decrypt the ciphertext and reveal the message in the process. This can also be done by the adversary, if it can find a vector of length n and also if it has the ability to correct k errors. Solutions to the problem in the Hamming metric have been presented using a set of indices [15] and its variants [146] to find the codeword with the smallest weight. Also, these solutions has been extended to the rank metric to guess the support that contains the error coordinates [10].

However, with available literature, no Post quantum based cryptosystem has been designed using codes associated with the Grassmannian under the Grassmann metric. Nevertheless, there is ample evidence that points to the fact there is a connection between the construction of a cryptosystem using a Grassmann based code and a Hamming based code. This is because of the link between the structure of these two codes as explained in this paper [147]. Also, no solution to the problem in the Grassmann metric has been proposed as regards to its use in cryptography. However, for coding applications, research on finding the minimum weight of codewords in the Grassmann metric has been proposed [148].

The Grassmannian can be divided into positive or negative depending whether the maximal minor of the generator matrix, in other words the determinant, is positive or negative. In other words, a negative Grassmannian has a negative minor while a positive Grassmannian has a positive minor. Furthermore, the positive Grassmannian has positive Plücker coordinates as well and the essence of using the positive Plücker coordinates as a solution to the syndrome decoding problem is to avoid changing the minor which would lead to erroneous results when swapping the columns of the generator matrix. Consequently, in the Grassmann metric, sets of Plücker coordinates are analogous to index sets used in the Hamming metric.

The Grassmannian defines a system of k-dimensional subspaces in an n -dimensional vector space of a finite field of Characteristic 2. It also includes a projection of n-k dimensional subspace that form unique pivot positions. These subspaces can be seen as vertices connected by edges, if and only if there is a trivial intersection between the subspaces. In the process a unit Grassmann distance is generated. Furthermore, for such a graph, sparse bi-adjacency matrix represents the nodes and the edges which can be decomposed into a set of positive Grassmannian Schubert cells [12]. These cells can be represented by a canonical matrix in a row echelon format with a leading one in each row. The missing element in each row can be modelled using Ferrer's diagram [149] which represents it as partitions. A relevant research question is this, are there codes associated to Grassmannian varieties with robust theoretical background that can be categorized as a sub family of Tanner graph codes? The synopsis to this question comes from the idea of using Grassmann support and its mathematical framework [10] on code based based cryptography in the rank metric. This parameter is usually used as a parameter for codes associated to Grassmann varieties. This inspires the thesis to connect the dot by expounding on the Grassmann support and its derivatives. Finally, in the theory of toric geometry [12], the planar graph that illustrates the totally non negative Grassmannian can be redesigned into a graph similar to a Tanner graph [150] and possessing the properties of such a graph. Consequently, Non-negative Grassmann codes is a graph based code that can be represented with vertices and nodes just like Tanner graph based codes. These nodes represent k dimensional subspaces and their intersections .

The solution of the syndrome decoding problem is generalized to the Grassmann metric by using Plücker coordinate. This is done by finding the subset of Plücker coordinate of codewords of minimum Grassmann weight and with minimum subspace distance. The Plücker coordinates of the totally positive Grassmnannian cells are the the columns of the Generator matrix of the code $C(k, n) \subset G_r(n, k)$ whose maximal minor is non-zero . As stated earlier, this matrix is a representation of an isomorphism. Such isomorphism has been studied in literature [151]. Families of codes associated to Grassmann varieties can be employed in the quantum safe code based cryptosystem because of its efficient decoding procedure [151] and probability to correct low weight codewords [149].

The adversary requires knowledge of the isomorphism in order to decompose the Generator matrix into its row echelon form. The boundary map would be

used as an isomorphism to map the k subset elements of the generator matrix into a point in the Grassmannian in order to find non-negative Plücker coordinates with minimum Grassmann distance. Furthermore, an a priori approach is proposed to find the low Grassmann weight vector by enumerating the basis based on a bound that is expressed as function of the number of positroid cells in the Grassmannian $Gr_{k,n}$ with weight k.

4.1.1 Contribution

The contribution of this chapter is to propose an alternative solution to the syndrome decoding problem in the Grassmann metric using Plücker coordinates. First, the theory of Plücker coordinates is extended with the transformation of planar graphs to non planar graph. Then, Gaussian decomposition is employed to reduce the generator matrix to form pivot columns of the parity check matrix. Thereafter, analytical bounds on the enumeration of the basis are presented together with numerical results on the failure probability and the cost of row operations. Finally, the shortest vector problem is generalized to the syndrome decoding problem and the probability mass distribution of the elements of the Grassmannian in their coordinate positions is analyzed using copula functions.

4.2 Preliminaries

4.2.1 Notation

In this section, a brief summary of some of the notation used in this chapter is provided. F_q represents finite field of q elements, F_{q^m} represents extension field of degree m, F_q^n represents vector spaces of dimension n over F_q, A represents $n \times m$ matrix, a represents a vector, $G_q(n)$ represents set of subspaces belonging to F_q^n(Grassmann), $E \oplus F$ represents smallest subspace ,$\langle A \rangle$ represents F_q span of A.

4.2.2 Coding Theory in the Rank Metric

Given a bijective mapping between a vector a and a matrix $A \in F_q^{m \times n}$, with a subspace of size $n-k$, the complexity of a combinatorics solution is given by $O(n-k)^3 m^3 q^{(n-k)\left[\frac{(k+1)m}{n}\right]} - m$ [10]. Lifting can be performed on an interleaved code by transforming the linear matrix code to a subspace by multiplying its transpose with an identity matrix. The linear matrix code $C[\, m \times n, k] \in F_{q^m}$ is a linear code generated by $(m \times n)$ matrices. The linear matrix code can be represented as a function of its basis by $C_j = \sum_{i=1}^{m} X_{ij}\beta_i \forall j \in \{1, \ldots, n\}$ where β_i is a basis

of a subspace F over F_{q^m}. The basis of a subspace over F_q multiplies C by a non zero element which does not affect the rank distance between codewords. The basis can also be a row of a generator matrix $G \in F_{q^m}^{k\times n}$ which has the complexity of $k(n-k)m^2 \log_2 q$ bits [8]. The dimension of the subspace determines the weight of the codeword and the number of subspaces is given by the Gaussian coefficient expressed as

$$\binom{n}{w}_q = \prod_{i=0}^{k-1} \frac{q^n - q^i}{q^w - q^i} \tag{4.1}$$

w is the weight and q^m and q^i are monomials over F_{q^m}.

In information set decoding, the probability of finding the codeword given a $[\ n, k, t+1]$ matrix code is given by

$$P_{dec} = \frac{\binom{n-k}{t}}{\binom{n}{t}} \tag{4.2}$$

with complexity $P_{dec} = O(1).2^{nH_2(t/n)-(1-k)H_2(t/(n-k))}$ where $H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ [16]. If the parity check matrix H is expressed with respect to $(n-k)\times n$ identity matrix, an $m \times k$ zero matrix and $(n-k-n)\times k$ random matrix code chosen uniformly as $H = (I/0/R)$ then the linear matrix code is called a simple code. The probability of decoding such a matrix when $m < \frac{m+n-\sqrt{(m-n)^2+4km}}{2}$ is given by $P_f \sim \frac{1}{q^{m-w+1}}$ as $q \to \infty$. The bound on the weight of the error vector x is given by the Gilbert-Varshamov bound [152] which is defined as

Definition 8. *The number of elements of a sphere S given integers n, m, q, t with radius $t \in F_{q^m}^n$ is equal to the number of spaces with $m \times n$ basis of dimension t. For $t \geq 1$, it follows that*

$$S = \prod_{j=0}^{t-1} \frac{(q^n - q^j)(q^m - q^j)}{q^t - q^j} \tag{4.3}$$

For a ball of radius t, the volume becomes $V = \sum_{i=0}^{t} S(i)$. Also, for a matrix code C, if $V \geq q^{m(n-k)}$ and $\sum_{j=0}^{d-2} \binom{n-1}{j} < 2^{n-k}$ then the smallest integer

t is referred to as the Gilbert-Varshamov bound.

4.2.3 Syndrome Decoding Problem

The Syndrome decoding problem is defined here in terms of complexity theory

Definition 9. *The a priori probability of finding a codeword x_i with non-zero codewords $\leq w$ given an integer which represent the ith column of an error that permutates Code C to C',such that $H^T x = s$, where $s \in_R F_{q^m}^{n-k}$ is a syndrome and H is a parity check matrix over F_{q^m}.*

Consequently, to generalize this problem to the Grassmannian metric, it has to be reduced to an instance of finding the Plücker coordinates of codewords with lowest Grassmann weight.

Definition 10. *Let Plücker coordinates be denoted as $\Delta_{I,J}(G) > 0$ which forms the columns of the generator matrix. The syndrome decoding problem is to find linearly dependent subset of Plücker coordinate with w columns such that $G_{i,j-k} \wedge v_j = u_i$ where a basis B is defined thus; $B = \{u_i, v_j | i \in I, j \in J\}$, a $k \times n-k$ matrix M_v and a $k \times n-k$ generator matrix G with rank k.*

4.2.4 Grasmmaninan theory

Definition 11. *Totally non-negative Grassmanninan [12] is the point in the Grassmannian with positive Plücker coordinates $\Delta_I \neq 0$*

In other words its maximal minor is positive and it can be combinatorially modelled into planar bipartite graph. The matroid of the totally positive Grassmannian is termed a positroid.

Definition 12. *The boundary measurement map [12] is defined as $b : R_{>0} \rightarrow G_{L_k}.A \in Gr_{n,k}$ where A is a $k \times n$ biadjacency matrix with a rank k which are represented by incoming boundary edges. The map depends on the coloring of the vertices.*

The matrix has a maximal minor $\Delta_I = 1$ that forms the Plücker coordinates on $Gr_{n,k}$ with column vectors $\frac{I}{A}$ that gives the basis of the subspace. Furthermore, the coordinates of A can be defined as follows with slight abuse of notation as $\varphi(A) = \langle (u_i + \sum_{j=1}^{n-k} A_{ij} v_j) \rangle \forall 1 \leq i \leq k$.

$R_{>0}$ is characterized by the set of all the biadjacency matrix A. The subspace in this set is a graph of a map from a projection to its orthornormal. In other words, $V \to V^{\perp}$. The direct sum expression is given by $V \oplus V^{\perp} \cong R^n$ with a basis $V = \{v_1, \ldots, v_a\}$.

Let the map of a subspace U to its local diffeomorphism be given as $\phi(u) = (\phi_1, \ldots, \phi_n)(u_1, \ldots, u_k)$, then it follows that the tangential space at any point of the map has a basis with coordinates $\{\frac{\partial \phi}{\partial u_1}, \ldots, \frac{\partial \phi}{\partial u_k}\}$. In other words, the tangential space can also be represented by the derivative of the Grassmann. If there is an open subspace in the Grassmannian $Gr_{n,k}$, then $U = \{W : W \cap V^{\perp} = \{0\}\} \subset R^k \times R^{n-k}$ for any $W \in U$.

There are complex numbers c_{ij} such that $v_i + \sum_{j=1}^{b} c_{ij} v_j \in W$ which is linearly isomorphic. This implies that $U(S) = \{v + Sv : v \in V\}$ such that $v \mapsto (V, S(v))$. If $v = 0$, then $U(C) = 0$ from the nullity of maps. If V is decomposed to subspaces P and Q where $Q \in U_A$ and U_A is a set of all subspace $P \subset V$ such that $V \cap U_A = \{0\}$, then there exist $P = (P \cap Q) \oplus P^{'})$ for some $P^{'}$ isomorphic to $P/(P \cap Q)$.

Furthermore, for a direct sum decomposition, the intersection of P and Q is trivial which now becomes $P + ((P \cap Q) \oplus P^{'}) = P \oplus Q^{'}$. If the subspace E is decomposed, there exist $E = (E \cap V) \oplus E^{'}$ for some $E^{'} \subset R^n$ where the intersection $E \cap V$ [151].

Finally, a function $F_k(V)$ is equivalent to the injective transformation $T : R^k \mapsto V$ and an open subset of $L(R^k, V)$ with a subspace of dimension $dim(F_k(V)) = kn$. In other words, $F_k(V)$ is the projective geometry of V and its quotient space generates the Grassmannian space.

Proposition 2. *Let V be a linear subspace and $V^{\perp}$ its orthonormal projection. Let U_A be a set of all projections $P_V \subset V$ through a map $U = v + Sv$. Then U_A lies in $L(V, E)$, if a linear isomorphism $T \in \pi^{-1}(U_A)$ exists.*

.

Proof. If there is an open subspace in the Grassmannian G_{n-k}, then the following holds

$$U = \{E \cap V^{\perp} = \{0\}\} \tag{4.4}$$

$$U(S) = \{v + Sv : v \in V\} : v \mapsto (v, S(v)) \tag{4.5}$$

where a subspace $S \subset V \oplus E$. This implies that $S \cap E = \{0\}$. Lets define two projections $P_{V'} : V^{'} \mapsto V$ and $P_V : V \mapsto V^{'}$ where $P_V(v)$ is related to $P_{V'}$ by the expression

$$P_V(v) = (P_{V'})^{-1}(v) - v \tag{4.6}$$

. Given U_A a set of all projections $P_V \subset V$, there is a linear isomorphism $T \in \pi^{-1}(U_A)$ and a projective geometry $F_K(v) = \pi^{-1}(U_A)$ where π^{-1} is an invertible function. Then it follows that the intersection of T and the biadjacency A is trivial that is $\pi(U_A \cap A = \{0\}$, if the function π can be inverted and if a map $f(T) = 0$. For $v \in V$, it is assumed that the k dimensional subspace is equivalent to its transformation for some $v^{'} \in V$ that is $v + S(v) = v^{'} + S^{'}(v^{'})$. It follows that

$$v - v^{'} = S^{'}(v^{'}) = S(v) \in E \cap V^{\perp} = \{0\}, \implies S(v) = S^{'}(v^{'}) \tag{4.7}$$

. Concatenating the linear isomorphism T with the projections $P_{V'}$ and P_V, results to

$$f_T(v) = (P_{V'} \circ T) \circ (P_V \circ T)^{-1} \forall v \in V \tag{4.8}$$

and if f restricts $S = S^{'}$ on $L(V, E)$ then it becomes

$$f_T : \pi^{-1}(U_A) \mapsto L(V, E) \implies P_V(v) = (P_{V'})^{-1}(v) - v = v + Sv \tag{4.9}$$

. This results to

$$P_V(v) = Sv \tag{4.10}$$

$$f_T(v) = (P_{V'} \circ T) \circ (P_V \circ T)^{-1} = id_{V,V^{\perp}} \tag{4.11}$$

□

4.3 Extending the theory on Non negative Grassmann

In this section, the totally non negative Grassmann would be illustrated using Tanner graph like constructions by transforming it from its planar structure to non-planar structure. This can be seen as intersecting the theory of distance transitive graph and coding theory based on the framework of Grassmann variety. First, the concept of boundary measurement maps is redefined and thereafter, a logical breakdown of how this map can be represented as a binary matrix is presented. The boundary measurement maps are designed as a mapping or transformation of vertex set in a planar bipartite graph to edge weights defined as a set of vertices in a cell in the Grassmannian graph. Given a set $I_f \subset I$, removing an element from the set, an embedding can be constructed from the bipartite to the Grassmannian as

$G_{r_{k,n}}(R) \to RP^{\binom{n}{k}-1}$ which forms a gauge transformations expressed as a function of matroids $Meas : R_{>0} \to G_{r_{k,n}}(R)$ where $G_{r_{k,n}}(R)$ is k planes on an n-dimensional space which is not affected by the ratios of $k \times k$ minors of a $k \times n$ code. Furthermore, an arbitrary edge function is selected such that $e : u \to v$ and if the vertex is coloured, another edge function is selected such that $e^{'} : v \to w$ by maximum revolution. Depending on the coloring, this maximum revolution can be clockwise or anticlockwise. This maximum revolution induces self intersections through the path. The boundary measurement can be defined as as $M_{ij} = \sum_{P:e\to e'}(-1)^{wind(R)}wt(P,y)$ where the factor $(-1)^{wind(R)}$ is bound by the number of connection between sources to the planar bipartite graph which is made up of n external nodes of perfect orientation and k sources of perfect orientation. $wt(P,y)$ is the weight of the path.

The planar bipartite graph structure with perfect orientation[11],[12] would be employed to buttress the idea. This is shown in Figure 4.1 and Figure 4.2. First, the planar bipartite graph is transformed into non-planar bipartite graph taking note of the sources and external nodes while labelling them accordingly for convenience purposes. If the row and column are of the same node, the code entry is set to 1, if there is no path connecting the nodes, the map code entry is set to 0. Finally, the condition in literature is modified to support the objective of the idea by stating that if there is a negative sign then the entry is set to 0 and set to 1 if otherwise. Consequently,a boundary measurement mapping A and B produces the Grassmannian $G_{r>0}(2,4)$ and $G_{r>0}(2,6)$ respectively which is constructed using the flows as regards to whether it is clockwise or anticlockwise as follows;

$$A = \begin{bmatrix} 1 & 0 & -t+x & -(y+xzt) \\ 0 & 1 & y & zt \end{bmatrix} \implies \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \to G_{r>0}(2,4) \tag{4.12}$$

The same procedure is extended to B as well

$$B = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \to G_{r>0}(2,6) \tag{4.13}$$

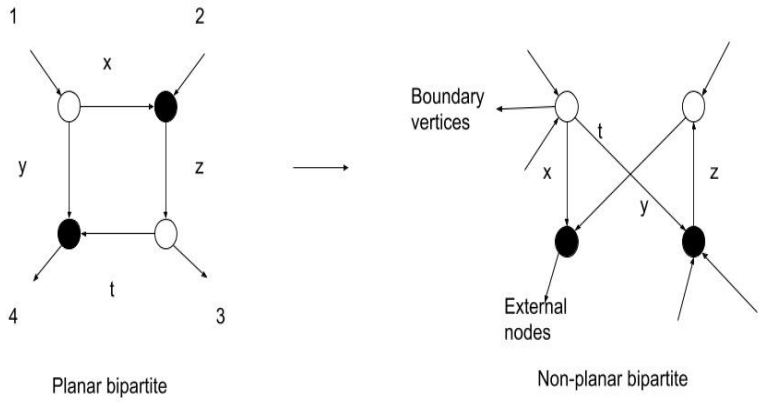

Figure 4.1: Non planar bipartite graph with perfect orientation containing 2 boundary vertices, 2 external nodes and a face transformed to its non planar structure

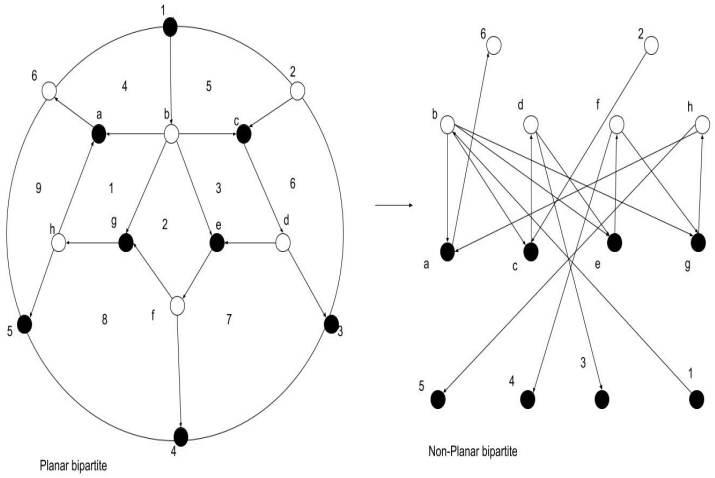

Figure 4.2: Non planar bipartite graph with perfect orientation containing 2 boundary vertices, 6 external nodes and 9 faces transformed to its non planar structure

The dimension of the Grassmanian parameterized from $G_{r>0}(2,4)$ is given as 4, then the number of boundary vertices k is computed as follows $k(n-k) = 4; k = 2$ while that of the Grassmannian parameterized $G_{r>0}(2,4)$ is given as 6, then the number of boundary vertices k is computed as follows $k(n-k) = 6; k = 2$

For a set $I = \{1,2\}$ and a minor $J = 2,6$,a modified Plücker coordinate for $\Delta_{2,6}$ can be computed as follows

$$\Delta_{26} = f/g = \frac{(1b + C2)(1b + ab)}{1 + C2} \tag{4.14}$$

4.4 Solution using Plücker coordinates

In this section, the concept of solving the syndrome decoding problem in the Grassmann metric using Plücker coordinates is presented.

Let $C \subset G_r^+(n,k) \in F_2^{k+l}$ be a code associated to the totally non-negative Grassmannian with a generator matrix $G \in F_2^{(k+l)\times l}$ and a subset of the matroid space Mat. Therefore, we have the matrix

$$G = \begin{pmatrix} g_0 & g_1 & g_2 & \cdots & g_n \\ g_0^q & g_1^q & g_2^q & \cdots & g_n^q \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ g_0^{q^{k-1}} & g_1^{q^{k-1}} & g_2^{q^{k-1}} & \cdots & g_n^{q^{k-1}} \end{pmatrix}.$$

The element of the Grassmannian are the F_q linear span of the columns of the generator matrix which contains the subspace $V = \langle g_i, \ldots, g_n^{q^{k-1}} \rangle \in R^k$ and the F_q linear span of the rows of the generator matrix contains the subspace $U = \langle g_i, \ldots, g_n \rangle \subset R^n$. The subspace V represents the isomorphism space while the subspace U represents the syndrome space. By employing Gaussian elimination and taking an instance of the boundary map $\tau \in b$, the code is lifted to $C^{'} = \tau(C)$ with generator matrix $G^{'}$ in row echelon form $G^{'} = \begin{pmatrix} I^l & O^l & H^{'} \\ O^{n-k-l} & I^{n-k-l} & H^{''} \end{pmatrix}$ where $H^{'} \in F_2^{(k+l)\times(k+l)}$,$H^{''} \in F_2^{(2k+l)\times(k+l)}$ are formed with pivot columns of dimension $k+l$ and $2k+l$ respectively. In addition, I^{n-k-l}, I^l are identity matrices of size $n-k-l$ and l respectively. $0^{n-k-l}, 0^l$ are zero matrices of size $n-k-l$ and l respectively.

The identity matrix I_u and the zero matrix O_V are both restricted to $n-k-l$ Plücker coordinate positions, $I_U = \begin{bmatrix} I^l \\ I^{n-k-l} \end{bmatrix}$ and $O_U = \begin{bmatrix} O^l \\ O^{n-k-l} \end{bmatrix}$. The matrices I_U and O_V are permutated by the parity check matrix H as follows $I_U H = \begin{bmatrix} H^{'} & I^l \\ H^{''} & I^{n-k-l} \end{bmatrix}$ and $O_V H = \begin{bmatrix} H^{'} & O^l \\ H^{''} & O^{n-k-l} \end{bmatrix}$. Furthermore permutation with error vector x to both matrices where x is generated by $k+l$ entries results to $I_U H x^T = \begin{bmatrix} H^{'} x^{'T} + x^{''T} \\ H^{''} x^{'T} + x^{''T} \end{bmatrix}$ and $O_V H x^T = \begin{bmatrix} H^{'} x^{'T} \\ H^{''} x^{'T} \end{bmatrix}$. Concatenating the matrices becomes

$$I_U O_V H x^T = \begin{bmatrix} (H^{'} x^{'T}.H^{'} x^{'T}) + (H^{'} x^{'T}.x^{''T}) \\ (H^{''} x^{'T}.H^{'} x^{'T}) + (H^{''} x^{'T}.x^{''T}) \end{bmatrix} \tag{4.15}$$

let $s = (s^{'}, s^{''})$ be the pivot of the syndrome then we have

$$I_U O_V s^T = \begin{bmatrix} H^{'} x^{'T} + O^l \\ H^{''} x^{'T} s^{'} + O^l \end{bmatrix} = \begin{bmatrix} H^{'} x^{'T} \\ H^{''} x^{'T} s^{'} \end{bmatrix}$$

Consequently, Plücker coordinates $\Delta_{I,J}(G)$ with size $k+l$ for $H^{'}$ are chosen and another Plücker coordinate $\Delta_{I,J}$ for $H^{''}$ where $I = \{i_1 <, \ldots, < i_k\}$ are k elements of G in monomial order. Thereafter, cycle shift to the columns of $H^{'}$ are implemented. Consequently,indices $i \in I$ is removed to form a basis of the subspace $V^{'} = \langle g_2, \ldots (-1)^{k-1} g_n^{q^{k-1}}, g_1 \rangle$. The process is repeated for the columns of $H^{'}$ to form the basis of the extended subspace $U^{'} = \langle g_2, \ldots, g_n, g_i \rangle$. A linear combination of the the $k-1$ columns of the subspace $V^{'}$ will form a vector $\tau(V^{'})$ and a linear combination of the n columns of the subspace $U^{'}$ will form a vector $\tau(\Delta_{U^{'}})$ with a pivot centered around $\tau \in b$. $\tau(V^{'}) + \tau(\Delta_{U^{'}})$ is added. Finally, the Grassmann weight is checked if it satisfies the distance criteria $d(V^{'} \cap U^{'}) \leq w - n + k - 1$ and the algorithm terminates if it does. if the condition is not satisfied, the process is repeated. It can be said that if the cyclic shift is applied, I becomes $I^{'}$. The Gaussian decomposition operation is a function of the ordering of the Plücker coordinate vectors.

Let $\Delta_{B(k,n)}$ be the Plücker coordinate of all subspaces with restriction in the first k Plücker coordinates $g_1, \ldots g_k^{q^{2k-n}}$. The $k \times k$ minor $\Delta_{B(n,k)}$ of the generator matrix $G^{'}$ is the set of k Plücker coordinates in $G_r^+(k,n)$. The instance of the boundary measurement map is validated by the adversary on the condition that $\Delta_{B(n,k)}(G) \neq 0$. It can be said that $B(k,n)$ which is the bounded affine permutations constitute the set of information sequences. The instance of the boundary measurement map can be represented by a Vandermonde matrix such that the Plücker coordinate is the column set of $I^{n-k-l} \in G^{'}$. Afterwards, the adversary selects an arbitrary subspace V with basis $V = \langle 0, v_1, \ldots v_{k+t} \rangle \subset C^{'}$ and choose the codewords with minimum weight $w \leq q^{\frac{k(k-1)}{2}}$. Finally, the Adversary checks if $d(U \cap V) \leq w$ and stops. By induction, it can be seen that there are $q^{\frac{k(k-1)}{2}} . \begin{bmatrix} k \\ r \end{bmatrix}_q$ ways of choosing the basis of the subspace V and $q^{\frac{k(k-1)}{2}} . \begin{bmatrix} n-r \\ k-r \end{bmatrix}_q$ ways of choosing subspace U.The proof of this claim is presented in Theorem 3. Therefore the probability of guessing

correctly the error free Plücker coordinates is given as $\frac{\begin{bmatrix} n-r \\ k-r \end{bmatrix}_q}{\begin{bmatrix} k \\ r \end{bmatrix}_q}$.

4.5 Bounds on Enumeration

In this section, bounds on the probability of enumeration of the basis of a lifted codeword $V^{'} \in V$ an element of the syndrome space and a lifted codeword $U^{'} \in U$ an element of the isomorphism that is a subset of the Grassmannian space in order to find the coordinates that indexes the elements of the vector with a minimum distance up to a Grassmann weight w are presented. In proposition 3, the bound on the Grassmann distance that is close to the distance of the error vector is derived, given subspaces U and V and a projection P_v that represents the isomorphism.

Proposition 3. *Let $U, V \in F_{q^m}$. As $q \mapsto 1$ and defining a map $P_v : F_q^n \mapsto F_q^{n-1}/V^{'}$ then $d(U,V) \leq 2q\begin{bmatrix} n \\ k \end{bmatrix}_q$*

Proof. Given k dimensional subspaces U, V of F_{q^m}, the Grassmann distance is given as $d(U,V) = k - dim(U \cap V)$. For vector spaces over the same field, there exist $dim(V \cap G) = dim(V) + dim(G) - dim(V \cup G)$. Therefore it follows that $d(U,V) = k - (dim(U) + dim(V) - dim(U \cup V)) \leq k - (k + k - (k - r) = r$ where r is the rank of the representative matrix. Given a subspace with dimension k, $\begin{bmatrix} n \\ k \end{bmatrix}_q = \prod_{i=0}^{k-1} \frac{q^n - q^i}{q^k - q^i}$, and selecting a $k-1$ dimensional subspace $V^{'}$ of F_q^{n-1}, the goal is to construct an arbitrary k dimensional subspace such that $V \cap V^{'} = \{0\}$. A basis $v \in V^{'}$, given as $v^{'} = \{v_1 < \dots v_{k-1}\} \subset N$ of a linear map defined as $P_v : F_q^n \to F_q^{n-1}/V^{'}$ is chosen to construct a bundle $\phi^{-1}(1) = V$. If $dimV^{'} = r$, then the number of bundles is equivalent to the number of enumerated basis of size $\{1, \dots, n-k\}$ over F_q which is q^{n-k}. This results in the identity

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \prod_{i=0}^{k-1} \frac{q^n - q^i}{q^k - q^i} = \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q \tag{4.16}$$

$$\tag{4.17}$$

It follows that for $0 < k < n$, it becomes

$$\leq \frac{q^{n-1}-1}{q^k-1} + q^{n-k}.\frac{q^{n-1}-1}{q^{k-1}-1} \leq \frac{q^{n-1}-1}{q^k-1} + \frac{(q^{n-k})(q^{n-1}-1)}{q^{k-1}-1} \tag{4.18}$$

Using a generalized identity [153] and doubling the right hand side of Equation (4.18), vectors except one of the q multiples of v can be computed as

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \sum_{i=0}^{k-1} q^{(n-k)(k-i)} \begin{bmatrix} n-i \\ i \end{bmatrix}_q \leq \tag{4.19}$$

$$\prod_{i=0}^{k-1} \frac{q^{n-i+1}-q}{q^i-1} \tag{4.20}$$

factorize q based on cardinality [154] Equation (4.12) becomes

$$\prod_{i=0}^{k-1} q\frac{q^{n-i}-1}{q^i-1} \tag{4.21}$$

□

Remark 1. *It can be seen that the error codewords can be formed by evaluating the monomials q^{n-1} with degree at most given in (4.21).*

Lemma 10. *The basis of the intersection of given subspaces U and V, induces a subset in $\Delta_I(G)$ that is linearly independent .*

Proof. The minimum distance of the total path of the subset is equivalent to the Plücker coordinate of the Grassmannian in projective space. Given a permutation $f_x(i) = min\{y \geq i/v_i \in span\{v_{i+1}, v_{i+2}, \dots v_j\}\}$ where v_i are the columns of the arbitrary space of S, taking basis $\{v_{i+1}, v_{i+2}, \dots v_j\}\}$ and extend it to $U \cap V$ as follows $v_{i+1}, v_{i+2}, \dots, v_j, e_{i-m+1}, \dots, e_i$ and $\{v_{i+1}, \dots, v_j, f_{i-m+1}, \dots, f_k\}$ through the path of the disk divided by a face $f \in U$ then, $P = \{e_{i-m+1}, f_{i-m+1}, \dots, e_i, f_i\}$ which forms a basis. The Plücker coordinate now becomes $\Delta_I(G) = \sum \prod_{P_i} wt(P_i)$, which implies that $\Delta_I(G)$ divides the vertex set Δ_I indexed by I an identity matrix such that each elements $e \in E$ and $f \in F$ induces a subset in $\Delta_I(G)$ □

It is necessary to use the analogy of section 4.3 to support the proof of Lemma 10.

Theorem 5. *The probability that k dimensional subspaces with a minimum distance d_{min} has $q^{i(i-1)/2}$ codewords in the Grassmannian G_r is bounded by $b_{r_k} \leq q^{i(i-1)/2} \begin{bmatrix} n \\ k \end{bmatrix}_q$.*

Proof. Intersecting k dimensional subspaces to $k+1$ dimensional subspaces with a minimum distance d_{min} will give the the isomorphism that maps of a point within a sphere that meets the Gilbert-Varshamov bound from the planar bipartite graph G to non-planar Grassmannian G_r, if $k+1 \in I$. For $k \notin I$ and with Plücker coordinate given as $\Delta_I(G) = \Delta_I(G_r) + w\Delta_I - \{k+1\} \cup \{k\}(G_r)$. This implies that $\Delta_{(I\ \{r\})\cup\{k\}} = (-1)^t b_{r_k} \geq 0$ where $t = |I \cap [\ w+1, k-1]\ |$. This results in the probability given as

$$(-1)^i \prod_{j=1}^{i} \frac{q^{j-1}qm - i + 1 - 1}{q^i - 1} = (-1)^i q^{i(i-1)/2} \begin{bmatrix} m \\ i \end{bmatrix} \tag{4.22}$$

□

Remark 2. *It can be seen from Theorem 5 and by induction that this probability depends on the the set that contains $m-i$ elements of the syndrome space containing codeword spanned by the basis of subspace V. Iterating over this set increases the probability that there are $q^{i(i-1)/2}$ neighbours in G_r for $0 \leq r \leq D$ where D is the maximum rank. Also, each row operation of the Gaussian elimination process preserves this probability.*

Given lifted codewords C_1 and C_2 with rank k_1 and k_2 respectively. C_1 and C_2 have subspaces V and U with basis $V = \{v_1, \ldots, v_{k_1}\}$ and $U = \{u_1, \ldots, u_{k_2}\}$. The the product of the spaces is bounded by $\langle VU \rangle \leq k_1 k_2$ where k_1 and k_2 are the dimensions of the support of V and U. If $k_1 k_2 < m$ then there exist a probability that the dimension of this support is $Pr(dim\langle VU \rangle) < k_1 k_2 \leq \frac{q^{k_1 k_2}}{q^m}$.

Corollary 1. *If the basis of a permutated codeword $V^{'}$ is random and the basis of a permutated codeword $U^{'}$ is fixed, then the probability that a projective space P_u and $V^{'}$ generates a random support with dimension k_1 is at least $1 - k_1 \frac{q^{k_1 k_2}}{q^m}$ where $dim\langle V^{'} P_u \rangle = k_1 k_2$.*

Proof. There exist a codeword $C \in P_u$ where P_u is a space and $C \notin F_q$. Given $\dim\langle V^{'} P_u^2 \rangle = k_1 k_2$ and a vector $x \in \langle U^{'} V^{'} \rangle$ with $x \notin V^{'}$ then the product CP_u is an element of the space P_u. □

Theorem 6. *Let $V^{'}$ be a subspace generated by a basis with dimension k_1 and $U^{'}$ is a subspace generated by a random basis such that the dimension $k_2^{'} = k_1^{'}(1 - k_2)$. If $V^{'} \cup \langle U^{'} V^{'} \rangle = \beta$ then its probability of enumeration is given as $1 - k_2 \frac{q^{2k_1 k_2^2 + k_2(k_2+1)}}{q^m}$.*

Proof. Let $\cap_i \beta_i^{-1} s = V^{'}$, then $V^{'} \cup \langle U^{'} V^{'} \rangle = \beta$ where $\langle U^{'} V^{'} \rangle$ is the product of the space with basis $V^{'}$ and $U^{'}$ and which generates a new basis β. If $V^{'}$ is random the dimension becomes $k_1^{'} k_2 - k_2 = k_2(k_1^{'} - 1)$. Therefore, a random space with a basis $U^{'}$ has a dimension $k_2^{'} = k_1^{'}(1 - k_2)$ as given. If $\langle U^{'} V^{'} \rangle \cap \langle U^{'} V^{'} \rangle_{-1} = V^{'}$ such that the dimension of a fixed basis $U^{'}$ is given as $\dim U^{'} = \dim(k_2) + U^{'} \beta^{-1}$, then it becomes equivalent to $\frac{k_2(k_2+1}{2} + U^{'} \beta^{-1}$. Multiplying both sides by 2 results to $k_2(k_2 + 1) + 2k_1 k_2^2$ with the given probability $\square$

Remark 3. *It can be seen from Theorem 6 that the probability scales with increase in the k_2 positions permutated by the enumerator.*

Theorem 7. *Given $U, V \in G_r(n, k)$ and $d(U, V) = dim(U) + dim(V) - 2dim(U \cap V) = k - r$ where k is the dimension of the subspace and r is the rank with integers l, p, m then the bound from the Gaussian coefficient on $d(U, V)$ given by*

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \sum_{k=0}^{\infty} \frac{q^{\frac{k(k-1)}{2}}}{(1-q)(1-q)^2 \dots (1-q^k)} . \begin{bmatrix} n-r \\ k-r \end{bmatrix}_q . \begin{bmatrix} r \\ k-m \end{bmatrix}_q . \begin{bmatrix} k \\ r \end{bmatrix}_q \tag{4.23}$$

Proof. Starting with a basis for U, $B_1 = (e_1, \dots, e_m)$, picking randomly linearly independent vector $x_{U_i} \in U$. Then a coordinate of x_{U_i} is accessed by permutation and replaced to produce a new basis for U after repetitive iterations to give $B_1 = e_1^{'}, \dots e_m^{'}, x_{U_1}, \dots x_{U_k}$. The count is updated as

$$\texttt{Count}_U = \prod_{k=0}^{U_i - 1} q^k = \sum_{k=0}^{U_i} q^{\frac{k-1}{2}} \binom{n}{k}_q . \tag{4.24}$$

The process repeats for subspace V with basis $B_2 = (f_1, \dots, f_m)$. Random linearly independent vectors $y_{V_i} \in V$ are selected and y_{V_i} is permuted for coordinates. This is replaced to produce a new basis for V after repetitive iterations to give $B_2 = f_1^{'}, \dots f_m^{'}, y_{V_1}, \dots y_{V_k}$. The count is updated to yield

$$\texttt{Count}_V = \prod_{k=0}^{V_i - 1} q^k - q^{k-r} = \sum_{k=0}^{V_i} q^{\frac{k(k-r)}{2}} \binom{k}{r}_q \tag{4.25}$$

Finally, starting with a basis for the intersection of the subspaces, $U \cap V$, $B_3 = (g_1, \dots, g_m)$, random linearly independent vector $z_i \in U \cap V$ are chosen to generate a new basis after repetitive iterations which are given as $B_{3'} = (g_1^{'}, \dots g_m^{'}, x_{U_1}, \dots, x_{U_k})$ and $B_{3''} = (g_1^{'}, \dots g_m^{'}, y_{V_1}, \dots, y_{V_k})$. Sampling an integer $l_i \in L$ where $L = \texttt{Vect}(x_U)$

and $p_i \in P$ where $P = \texttt{Vect}(y_V)$ and updating the count results to

$$\texttt{Count}_* = \prod_{k=0}^{U_i - V_i - 1} q^k - q^{k-r+t} - q^{k-r+p} = \sum_{k=0}^{U_i - V_i - 1} q^{\frac{k(k-r)}{2}} \begin{bmatrix} n-r \\ k-r \end{bmatrix}_q . \tag{4.26}$$

$$\begin{bmatrix} r \\ k-t \end{bmatrix}_q . \begin{bmatrix} k \\ r \end{bmatrix}_q$$

The bound can be computed from the total of the counts as, $\texttt{Count} = \texttt{Count}_U + \texttt{Count}_V + \texttt{Count}_*$. It follows that $U = \texttt{span}\{g_i, x_{U_i}\}$, $V = span\{g_i, y_{V_i}\}$ and $U \cap V = \texttt{span}\{g_i\}$ □

Corollary 2. *Let the linear parity check equation of a lifted Grassmann code with basis α_{ij} be $H^{'} x^T = \sum_{l=1}^{n} \sum_{j=1}^{k} \alpha_{ij} H_l^{'} V_j = 0$, there exist a probability*

$$Pr(U \cap V) = \frac{q^{\frac{k-1}{2}} \binom{n}{k}_q}{q^{\frac{k(k-r)}{2}} \binom{k}{r}_q} \propto q^{\frac{k(k-r)}{2}(n-k)} \tag{4.27}$$

with complexity $O(\frac{(n-k)^2}{2} q^{\frac{k(k-r)}{2}(n-k)})$.

Theorem 8. *if the dimension of the vector space $\forall d \leq 2$, then the number of monomials to be evaluated by the enumerator is given by $\sum_{\alpha=1}^{d} \binom{n}{l} \binom{\alpha}{n}^d \left(1 - \frac{\alpha}{n}\right)^{n-l} x^d$.*

Theorem 8 gives a closed form expression for the average number of iterations and the proof is given in Appendix C.

4.6 Experiment on failure probability and cost of enumeration

In this section, we test the failure probability of enumerating these basis and compare the results with code that employs the hamming metric. It is also important the algorithm runs with as much sets as possible to make the iteration process smooth and efficient. Data collection through simulations had an impact on the memory of the Computer used. In these experiments, an AMD Ryzen 3 2200U laptop was used with Radeon Vega Mobile Gfx graphic card. The processor speed of the computer was 2500MHz with 2 cores and 4 logical processors. The

clock speed was 2.5GHz. Due to the limitation of the memory, the experiments were conducted with little amount of code sizes. However, these experiments can be scaled up without much impact on the fidelity of the result obtained.

4.6.1 Probability of failure

In this section, the results of experiments on the probability of failure while finding codewords of minimum distance are presented. The implementation [155] was optimized for this purpose. Theoretical analysis on the comparison between codes from the Grassmann and Hamming metric has been studied [147]. This thesis goes further by experimentally analysing the implication of this comparison on the security of a code based cryptosystem. The importance of this property on semantic security based on Indistinguishability using a Chosen ciphertext attack cannot be overemphasized. This is due to the presence of negligible error patterns in the received word. The lower this probability, the higher chance of the quantum adversary to distinguish between random instances of the ciphertext. In this experiment, the number of coordinate sets is given as 2^l for each level of security under investigation where l is the size of the set. For 128-bit security level, it becomes $32,768$ and the result is shown in Figure 4.3. For 256-bit, the number of coordinate sets becomes 1048576 and the result is shown in Figure 4.4. For 512-security level, it becomes $33,554,432$ and the result is shown in Figure 4.5. Finally, for 1024-security level, it becomes 1073741824 and the result is shown in Figure 4.6. The standard deviation of the distribution σ for all security levels is varied from 0.30 to 0.85 for cryptography purposes. To compute the amount of Gaussian elimination operation carried out, the formula $\frac{1}{2}(n-k)k^2$ is used. This is shown in Table 4.1. This formula relates the number of coordinate sets to the Gaussian decomposition operations. It can be seen from Table 4.1, that the number of rows to be reduced increases as the security level increases. This is due to size of the coordinate set for each security level which is bounded by $n-k$. The reason for this, is to limit the frequent failure of the algorithm due to the probabilistic approach of permutating the columns. However, this comes at a great computational cost. From the result shown, It can be seen that the failure probability of the Non-negative Grassmannian code is smaller than the failure probability of the LDPC code. The implication of this is that the Non-negative Grassmannian code based cryptosystem is more secured than the LDPC code based cryptosystem under the IND-CPA(Indistinguishability under the Chosen Plaintext Attack) model. This is because in the IND-CPA model, the probability error must

be negligible in order for the probability polynomial adversary to find it hard to be able to distinguish a message symbol sampled from a theoretical distribution from that sampled from an arbitrary distribution. In Figure 4.3, at a standard deviation of 0.50, the failure probability of the Non-negative Grassmann code is less than that of the LDPC code by 1.18 percent, In Figure 4.4, at a standard deviation of 0.50, the failure probability of the Non-negative Grassmann code is less than that of the LDPC code by 3.23 percent. In Figure 4.5, at a standard deviation of 0.50, the failure probability of the Non-negative Grassmann code is less that of the LDPC code by 2.34 percent. Finally in Figure 4.6 at a standard deviation of 0.50, the failure probability of the Non-negative Grassmann code is less than that of the LDPC code by 3.17 percent. As the security level increases, the probability that k subspace has $q^{i(i-1)/2}$ connected subspaces increases which induces some level of randomness on the choice of Plücker coordinates and in the process expanding the probability that a zero error pattern is contained in the syndrome space. This can be seen in the reduction in the error floor as the security level increases.

Table 4.1: Row reduction operations as a function of Security level

Security level	Row reduction
128	131072
256	1048576
512	8388608
1024	67108864

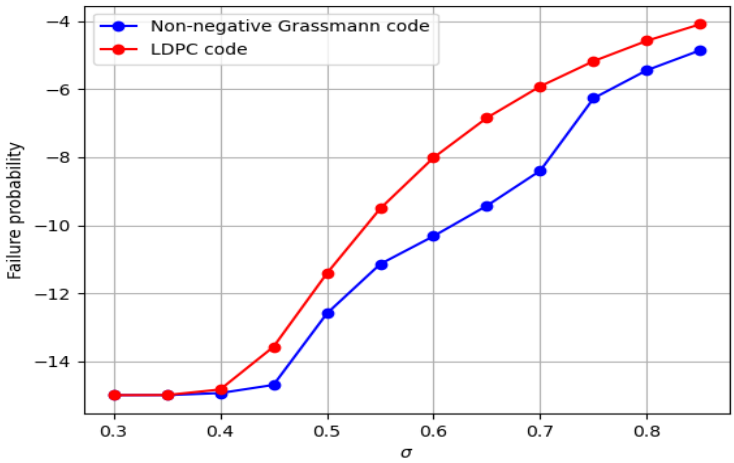

Figure 4.3: Probability of failure for 128-bit security, security parametr $l = 15$

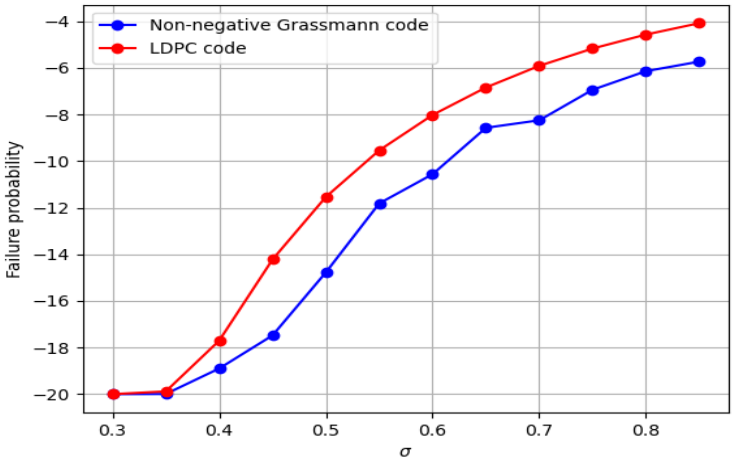

Figure 4.4: Probability of failure for 256-bit security,security parameter $l = 20$

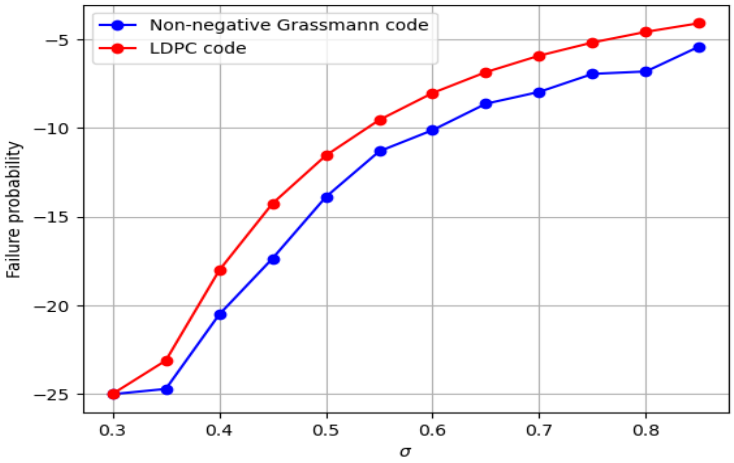

Figure 4.5: Probability of failure for 256-bit security,security parameter$l = 25$

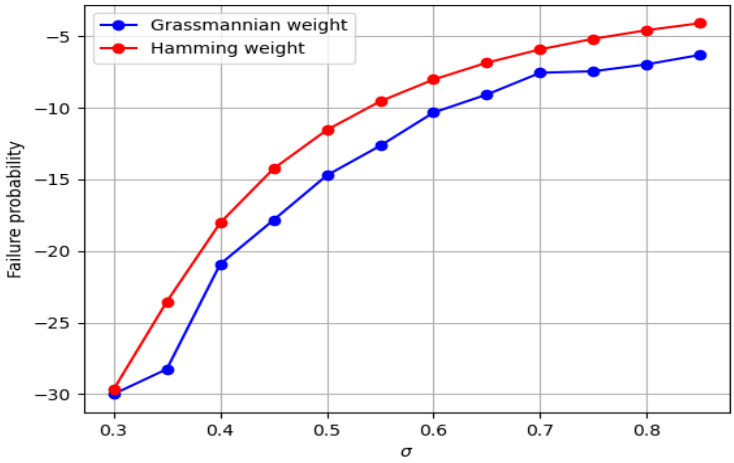

Figure 4.6: Probability of failure for 1024-bit security, security parameter $l = 30$

4.6.2 Cost of Enumeration

The experiment continues in this section were the implementation [156] was optimized to test the cost of iterating over the rows of the code with increase in code length. The results are presented in Figure. 4.7 for finite field of characteristic 2 and in Figure 4.8 for a finite field of characteristic 2 and extension 2. The result shows that the cost of iterating over rows of the Non negative Grassmann code is higher than of the LDPC code with increasing code length. At a code length of $n = 100$, the cost of row operations is higher by 5.81 percent. This shows that Non negative Grassmann code based cryptosystem is stronger against ISD attack than LDPC code. This is good for Post quantum security. In Figure 4.8, the field size was extended by 2 and a difference of 29.4 percent was recorded. The huge difference is a result of the large size of the coefficients of the polynomial linear equations with variable q, the field size which in turn increases the size of the basis of $k + 1$ subspaces of dimension $n = 1$.

Quantum security can be obtained by dividing the security bits by 2. This implies that for 128 bit security, the equivalent quantum security is 56 bits and in order to make the density of the syndrome close to 1, the parameters must satisfy the conditions specified by Bernstein et al [157]. The data size and computational time are linear in $logq$ while the complexity of combinatorics are polynomial on q making it difficult to break the encryption key. The decoding error with failure probability is equivalent $\frac{1}{q^{l'-2wr+1}}$ [158] and the key size increase inversely to an increase in the

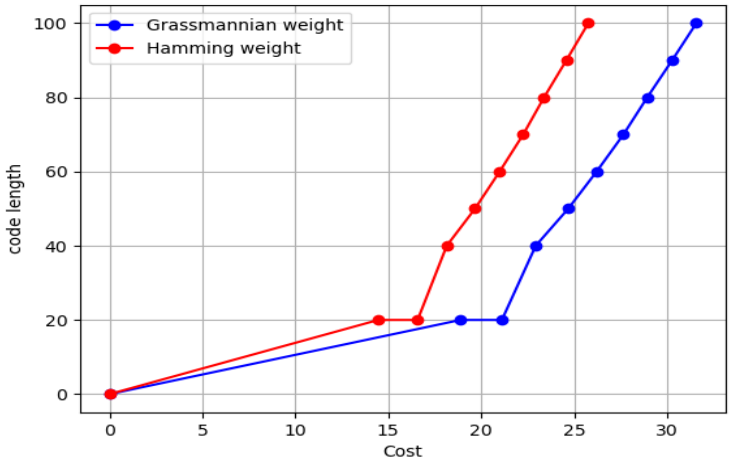

Figure 4.7: Cost of row ISD operations,field size $q = 2$

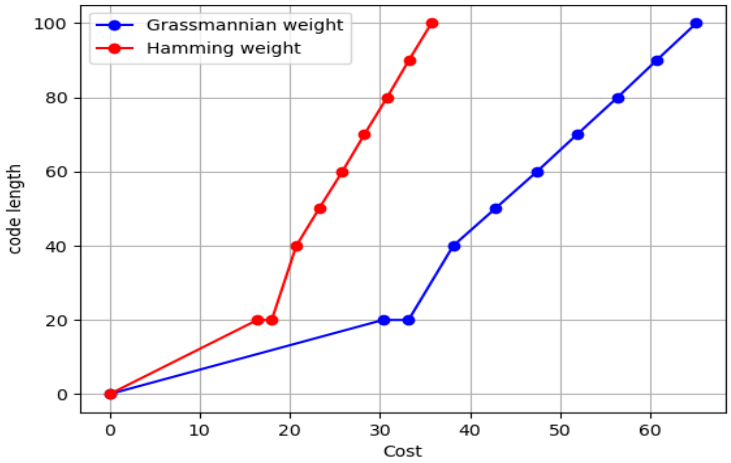

Figure 4.8: Cost of ISD row operations,field size $q = 2^2$

probability of the decoding failure. In the presence of cyclic vectors, classical attacks makes it possible to obtain the Plücker coordinates of the permutated codewords. In Table 4.2, proposed parameters are presented for code in the rank metric where n is the code length, k is the code dimension, m is the degree of extension field, q is the prime, w is the error weight which is compared to other parameters from related works. The works compared in the table were variants of ISD employed in cryptanalyzing Code based cryptography in the rank metric. From the complexity derived from Theorem 3 and Remark 2, it can be deduced that the complexity of the ISD decomposition on the input of the proposed parameters is 2^{23} which is below the claimed security level of 2^{128}. This proves that the complexity of using the Plücker coordinates depends on the minimum distance as derived from proposition 2.

Table 4.2: Comparison with Parameters in the Rank metric

n	**k**	**m**	**q**	**w**	**Security**	
67	7	89	2	5	128	[159]
100	80	96	2	5	192	[158]
100	80	96	2	5	192	[69]
67	22	71	2	11	133	[160]
110	7	18	2	12	128	proposed parameters

4.7 Generalizing from the Shortest Vector Problem

The syndrome decoding problem can be transformed to an instance of the shortest vector problem. In this case, the shortest vector is the vector with the smallest weight that produces a syndrome in a code. The enumeration process of finding this vector is transformed to decoding the vector using the probability distribution of the vector. It is assumed the vector is transmitted through a communication channel. In this section, the probability distribution of the coordinate positions are modelled using copula functions. This analogous to the weight distribution property used in Information set decoding to analyze different coordinate positions in order to find the minimum weight vector. The assumption is that the vector with the minimum weight vis a vis distance must also have the lowest probability of being decoded for all instances. Consequently, the probability of failure results obtained in the previous section depends on the probability distribution of the random vectors of the generator matrix of the code. There is the possibility of linking the theory of probability with the weight finding problem of the syndrome decoding problem

which would be a focal point for future study. This stems from the dependency between coordinate positions with various weight distributions. Here, the probability of either failure or success depends on the probability distribution of the columns with the weight distribution of the vector that solves the problem. The elements of the Non negative is assumed to be dependent and also sampled from a Gaussian distribution. In order to estimate the distribution of the coordinate positions, copulas are employed. The disadvantage of using Markov chains is because the sum over all transitional probabilities is not adequate to estimate the dependence of the positroid cells of the Non-negative Grassmann code. However, in experimentation, Markov chain are used for comparison. From research carried out, the probability distribution of graph based codes cannot be neglected in analyzing its usefulness for cryptographic purposes. The tail dependence coefficient is uniformly distributed and if the Legendre measure $\lambda_u = 0$, then the component are independent. The partial derivatives of the marginal cumulative distribution acts as a function of the elements of the transformation. This results in the product of the copula density and the marginal probability density function of the positive Grassmannian cells given as

$$\frac{\partial F(x_1, \ldots, x_n)}{\partial(u_1, \ldots, u_n)} = \frac{\partial^n C(x_1, \ldots, x_n)}{\partial(u_1, \ldots, u_n)} \times \prod_{i=1}^{n} f_i(x_i) \tag{4.28}$$

Families of copulas are defined based on the underlying distribution vis a vis standard multivariate normal distribution. This is defined by a correlation matrix R_G denoted as follows $C(\phi_{RG}(x_1), \ldots, \phi_{RG}(x_n))$ and given by

$$C(\phi(x_1), \ldots, \phi(x_n)) = \frac{\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\Sigma|}} \exp(\frac{-x^T x}{2\Sigma}))}{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x_i^2}{2})} = \frac{1}{\sqrt{|\Sigma|}} \exp(-\frac{1}{2} x^T (R^{-1} - I)x) \tag{4.29}$$

Given 2 distributions, the divergence can be modelled using Kullback Leiber divergence as [17]

$$KLD(f(x;\phi_1)||g(x;\phi_2)) = \sum_{i=1}^{d} KLD(f_i(x_i;\phi_{i,1})||g_i(x_i;\phi_{i,2})) + 0.5(\texttt{Trace}(\Sigma_2^{-1}\Sigma_1 + log\frac{|\Sigma_2|}{|\Sigma_1|} - d) \tag{4.30}$$

The properties of copula can be defined as follows; if there exists $k \in \{1, \ldots, n\}$ such that the component of the transformation at the kth position is equal to zero, then $C(u) = 0$. This is for all component of the transformation sampled from a uniform distribution. Also, if all the components of the transformation are approximated as unity, except the kth position, then $C(u) = u_k$. Finally if each component of the transformation in the kth position is increasing, then $C(u)$ increases by an equal amount. Furthermore, the conditional posterior distribution of the marginal of the syndrome space in the rank metric is given as $f_{x/y}(x_1, \ldots, x_n) = \prod_{i=1}^{m} f_{y_i/x} \times \prod_{i=1}^{n} f_{x_i/x_i}$ where $H_{ij} \neq 0, \forall E = 0$ and $\prod_{i=1}^{n} f_{x_i/x_i}$ is the independent realization. By exploiting the dependency between subspaces as shown in Figure 4.9, copula function can be applied to approximate the joint distribution of the coordinate positions. The following lemma supports this idea;

Lemma 11. *Given quotient space of the projective geometry as P_v, $F_K(v)$ is orthogonal to the support of the copula* $\mathtt{sup}(C)$*, if there exists a Legendre measure λ_n that is stochastic based on the magnitude of the copula function.*

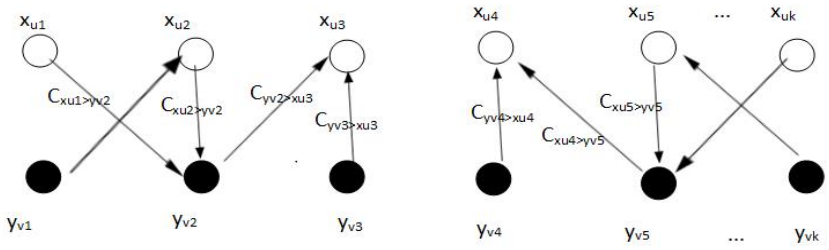

Figure 4.9: Dependency modeling of subspaces given a basis with span $\langle x_U \rangle$ and $\langle y_V \rangle$ using Copula functions

Proof. Given a subset, $F_K(v) \subseteq R^{n+m}$ such that $\lambda_n(P_V(G \cap F)) = 0$, for some subspace F, then we have $P_V(G \cap F) \supset P_{V'}(G \cap F) = \{0\}$. $\exists$ a subspace $V \subseteq R^n$ such that $P_V(V \cap F) \subseteq U$ where U parameterizes the Grassmannian space. Consequently, the copula is defined as the function that characterizes the dependency

as $C = C_{U,F}$. This implies that the volume that spans the copula, $Vol(C) = P(U \in P_V(F_K(v) \cap F) = P((U,F)) \in F_K(v)$ where $\lambda_n << Vol(C)$, if $n \geq 2$. By definition, $\texttt{supp}(C) \subseteq F$ [161] which follows that for a subspace E, $\lambda_n(P_V(V \cap F)) > 0$. This implies that $F_K(v) \cap \texttt{supp}C \neq 0$ □

From the above proof, it can be concluded that if $F_K(v) \in U_A$ such that $\lambda(P_{V'}(F_K(v)) \cap \texttt{supp}(C)) = 0$ is satisfied, then for a random point of the Grassmannian, $x \notin \texttt{supp}(C)$. The dependency property states the intersection cannot be a subset of the $k-1$ dimensional projection orthornormal to the subspace V.

The conditional distribution would be utilized as a smoothing function to find the coordinate with zero error positions.

$$\zeta(x^{(k)}) = E_{y/x,\hat{x}} \ln f(x;\hat{x}, P(y/x) \ = -\frac{n}{2}\ln|\hat{x}|\{\texttt{Tr}[R_G^{-1} - I] + \tag{4.31}$$
$$P(y)\int \exp \texttt{Trace}\frac{(-\frac{1}{2}y^T x(S_{NR}I - vx(1-\pi))y)^{-1}}{-\frac{1}{2}y^T y + \frac{1}{2}y^T x x^T y + ||\pi||^2} d\pi(x)\}$$

and update on the kth position is given by

$$\hat{x}^{(k+1)} = x^{(k)} + \Delta_k(\int f(x;\hat{x})d\pi(x)) \tag{4.32}$$

if the corresponding message is proportional to the conditional probability defined, each update in the iteration process leads to a better approximation of the coordinate with zero error positions. If $\hat{x}^{k+1} \geq x^{k+1}$ the copula monotonically reaches a point of convergence to its global maxima.

4.7.1 Distribution estimation

The approach in this paper by Gohary [13] was adapted to approximate the conditional distribution. This was done by estimating the reliable positions that maximizes the probability of estimating the codeword through splitting the generator matrix . By splitting the matrix, the upper triangular structure contains the information that enables the location of coordinate positions of minimum distance. With an arbitrary reference point, the distance between the positions that maximizes the conditional distribution is computed. The probability of avoiding decoding failure is if the codewords with minimum distance is within a bound $\hat{Q}_i - \hat{Q}_j$ where $\hat{Q}_i$ and $\hat{Q}_j$ are the chosen threshold values. By employing Cauchy Schwartz inequality, the probability that the codewords is within the subspace can be computed as follows;

Given a set of codewords $\{X_i\}_{i=1}$ generated by $h_0 = \hat{E}S^{-\frac{1}{2}}\omega_o$, then we have

$$||\omega_l - x_l|| = ||\omega_l - \hat{E}S^{-\frac{1}{2}}\omega_l|| \leq ||(I - \hat{E}S^{-\frac{1}{2}}||||\omega_l|| \leq \tag{4.33}$$

$$\texttt{max}\{|1 - \frac{\hat{E}}{\hat{Q}_i}|, |1 - \frac{\hat{E}}{\hat{Q}_j}\}$$

It is proposed in this thesis that the optimal values of $\hat{Q}_i$ and $\hat{Q}_j$ contain the codewords that maximizes the conditional distribution. The conditional distribution of y with respect to x as a function of projection π_T is given as follows

$$P(y/x) = P(y) \int \exp \texttt{Tr} \frac{(-\frac{1}{2}y^T x x^T y (S_{NR}{}^{-1} d\pi(x))}{-\frac{1}{2}(y-x)^T(y-x) + ||\pi||^2} = \tag{4.34}$$

$$P(y) \int \exp \texttt{Tr} \frac{(-\frac{1}{2}y^T x (S_{NR} I - vx(1-\pi))y)^{-1}}{-\frac{1}{2}y^T y + \frac{1}{2}y^T x x^T y + ||\pi||^2}$$

The codeword that maximizes the probability of decoding success can be found by solving the equation

$$\hat{x} = \texttt{arg min Tr}(-\frac{1}{2}y^T x x^T y + ||\pi||^2) \tag{4.35}$$

where $\pi = H(H^{'} S_{NR}^{-1} H)y$ is the projection onto the column space of y^T

Lemma 12. *Given the decomposition of the received codeword $y = Q_y R$, the correlation R is given by $(X\lambda X^T + \mu_r I)$ if $I(X, R) = 0$ and $I(H, Q_y) = 0$,*

Proof. It follows by induction that

$$R = X[\mu I + \lambda]^{-1} - \hat{G}(\lambda)^{-1} X^T + \lambda)^{-\frac{1}{2}} \tag{4.36}$$

$$= X\mu_r I + \frac{X\lambda}{G\lambda} - \sqrt{X^T + \lambda} =$$

$$[X\mu I + X(G)^{-1} - X^T] + \lambda = \varphi + \lambda.$$

where $\hat{G}$ is the orthonormal basis for R while λ is signal eigenvalue and φ is the projection. □

The projection onto column space defines a spherical ball with a radius characterizing the projective distance. This can also be generalized to be the minimum distance between the reference point and the codeword close to it. This codeword maximizes the distribution which can correct up to $\mu - 1$ errors where μ is the radius. If the radius is large, the decoding failure occurs which results in finding the codeword with the minimum distance inefficient.

Lemma 13. *The probability of finding the coordinate positions for a received codeword with minimum distance based on the expectation of the BIAWGN channel,*

if the nth central moment is about the mean of the distribution is given by $\mu_n = \frac{x-\mu}{4\mu}.e^{\frac{-(x-\mu)^2}{\sqrt{4\mu}}}dx$

Lemma 14. *Given the syndrome, the probability of generating an output vertex is bounded by the discrete analog of the derivative of the dimension of the vector space* $\forall d \geq 2$.

The proof of lemma 13 and lemma 14 is found in Appendices A and B respectively.

Theorem 9. *Given the marginals of the basis that span the subspaces* U *and* V *as* $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}$ *and the correlation coefficient* $\rho, 0 < \rho < 1$ *sampled from a Gaussian distribution in unit cube* $\begin{bmatrix} 0 & 1 \end{bmatrix}$*, the bound on the probability density function of the syndrome for the codewords is given by*

$$\int_{-\infty}^{\infty} N(\rho x, 1-\rho^2)dx = \frac{1}{\sqrt{1-\rho^2}} \exp\left(\frac{-\rho^2 x_1^2 + 2\rho x_2 x}{2(1-\rho^2)}\right). \tag{4.37}$$

The proof is found in Appendix D.

4.8 Experiment

Bit error rate performance experiment was carried out to simulate the probability distribution of the codewords of the Non negative Grassmann code based on coordinate positions using copula functions. The probability distribution of LDPC code using Markov chain for index sets and the probability distribution of fountain codes using marginals from belief propagation was also simulated. The simulations were carried out over 10000 Monte Carlo iterations. A bi-adjacency matrix of size 32 and dimension 64 was used. The Constellation size M was increased in steps of 5 to a maximum of 20. From the results, it could be seen that increasing the size improves the error correction capability. Furthermore,it can be observed from the plots that using marginals from belief propagation [162] has a slightly better BER performance than the Copula based simulation. In Figure 4.10, at a bit error rate 10^{-3}, the BER of using marginals performs at 0.6dB better than that of using copula. However, using copula performs within 0.4dB better than using index set based Markov chains. This is due to the complexity of modeling the dependency between the subspaces using the copula function. This also means that the subspaces might not be dependent, however further methods of improving the performance based on the dependence criteria might be considered.

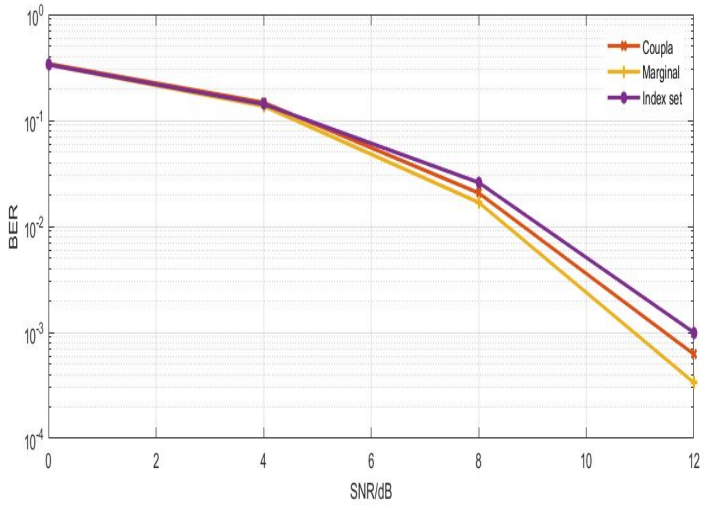

Figure 4.10: BER performance with constellation size M=5

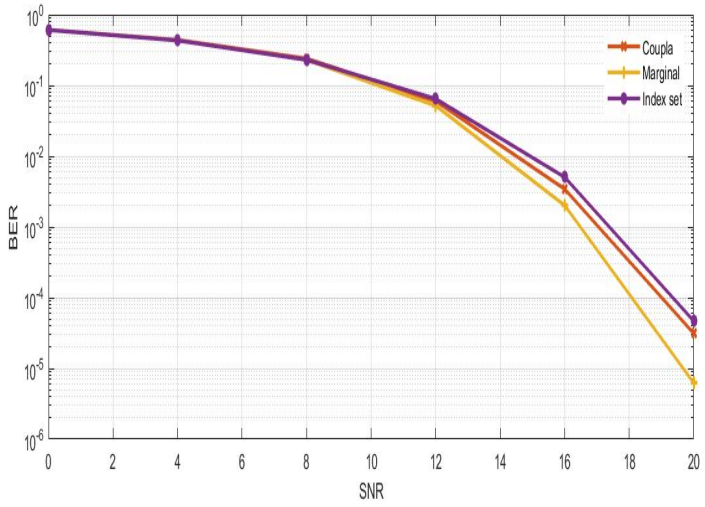

Figure 4.11: BER performance with constellation size M=10

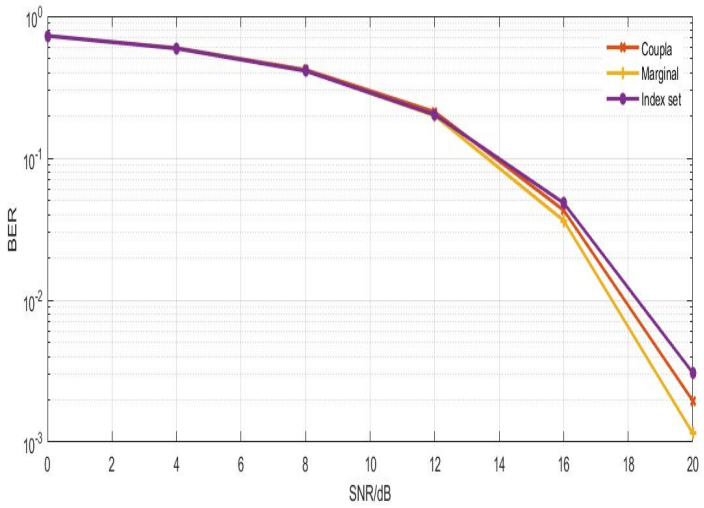

Figure 4.12: BER performance with constellation size M=15

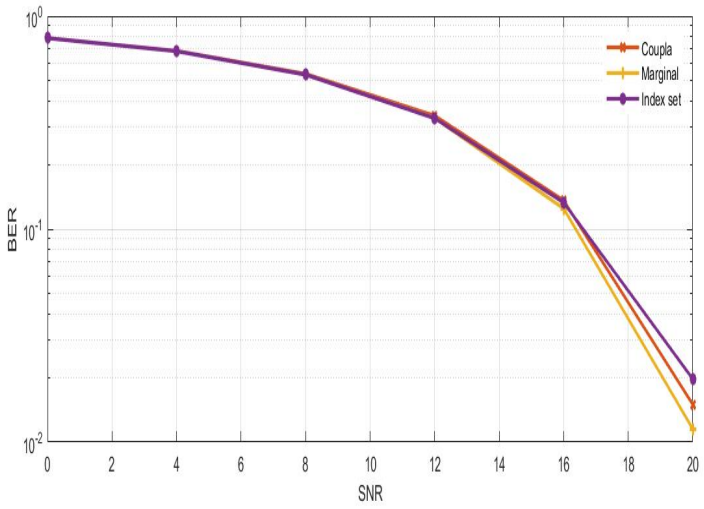

Figure 4.13: BER performance with constellation size M=20

In the first subset of coordinates with size $k \times k$, the Bit error rate value is directly proportional to the p columns of the lifted Non negative Grassmann code . As the weight of the columns increases, the bit error rate increases. It can be seen that the weight of the Non Grassmann code is higher than that of the fountain code but slightly less than that of LDPC code. This is due to the presence of cycles in the LDPC code. This weight translates to the weight of the vectors. In order words, finding the smallest weight codeword depend on the bit error rate. The lower the

bit error rate, the higher the probability of finding the lowest codeword and solving the syndrome decoding problem.

4.9 Chapter Summary

The syndrome decoding problem as a computationally hard primitive has been used in code based cryptosystem to secure information systems from quantum based solutions. In this chapter, solution to the problem using Plücker set is generalized to the Grassmann metric for codes associated with the totally non negative Grassmanninan. To visualize the rows and columns of the parity matrix of the totally non negative Grassmannian, a planar framework is employed. This planar structure can be transformed to non planar structure which is akin to a Tanner graph. The Boundary measurement mapping can be seen as gauge transformations expressed as a function of matroids. To form this image, edge weights are selected at random and grouped using a colouring function. The bounds on probability of finding the vector coordinates with the smallest Grassmann weight using exhaustive search of the basis by enumeration was presented. Consequently,the Generator matrix was decomposed into positroid cells using Gaussian elimination to find linearly dependent subsets of the Pl¨ukcer coordinates with minimal non-zero coordinates and in which the the maximal minor is totally positive. Finally, numerical results presented showed that the Non negative Grassmann code had a low probability of failure when compared with an LDPC code. This implies that the error floor of the LDPC code is higher than that of the Non-negative Grassmann code. Also, for increase in the code length, the decoding cost for the totally non negative Grassmann code was higher than the LDPC code. This validates the notion that the Non negative Grassmann code in the Grassmann metric was more secure under the Chosen ciphertext model when compared to the LDPC code in the Hamming metric. Due to its robust security credentials, this code is recommended to construct future post quantum encryption schemes.

Chapter 5

Isomorphism of Polynomial problem : Solution using New Mersenne Number transform

5.1 Introduction

Isomorphism is defined as a vector space mapping that preserves the unique mathematical properties of the space. A multivariate quadratic polynomial is defined as a system of univariate equations and variables that lie in a base field or an extension field. Isomorphism of polynomial problem is a computational hard assumption employed in multivariate polynomial cryptography[23]. The computational hardness assumption of Multivariate polynomial cryptography lies in the hardness of solving random non-linear multivariate quadratic equations over finite fields. An instance of the multivariate encryption scheme using Hidden Field equation has been developed which employs a small field of characteristics 2 and a degree extension m[22]. These schemes involves translating the multivariate polynomial to a univariate polynomial over the degree m extension field in order to generate a public key. The core quadratic map that makes up the public key is masked by two invertible affine transformations over a finite field[22]. The system of equations in n variables that make up the public key is constructed using functional composition of the affine transformation with the core map. However, such polynomial systems can be easily decomposed if the degree d of the core map is known[163]. Due to its vulnerability to attacks, variants of the HFE have been developed which employ layers of vinegar equations[88]. Unfortunately, the map of the variants suffers from

the limitations of bijectiveness. This is because of the randomness of the affine transformations which is bounded by the degree of the extension field. Due to the bijective property of the map, it becomes difficult to invert. An injective map with large codomain have been proposed as an alternative leading to efficient masking and inversion . Techniques have been proposed to increase the codomain size of the public key using specialized matrices with a large algebraic structure. This makes it simple to hide the structure of the injective map[40]. Key recovery attacks exploit the rank of the representative matrix of the quadratic form that make up the core map. They also exploit the ease of inverting the injective affine map. The disadvantage of using affine transformation as a trapdoor that preserves the isomorphism of polynomial problem is its ease of inversion (which makes the core map easy to attack) and its density (which leads to increase in key size) [164].

In this chapter, an isomorphism is constructed using the orthogonal kernel function of the New Mersenne Number transform[20] to hide the structure of the core map. This process works efficiently for core maps from monic polynomial sequence with rational coefficients. Using functional composition of the New Mersenne Number transform with the core map, the public key can be generated. Using the New Mersenne Number Transform provides an efficient method of evaluating the n coordinates of the core map which is the secret key of the multivariate cryptosystem. It can be said that the multivariate polynomial is orthogonal if and only if the polynomial is isomorphic. This is because the rank of the kernel matrix is equal to the rank of its transpose. Also the kernel matrix is of full rank and is equal to the dimension of the vector space. However, the reverse of the statement is not always the case. The security of the isomorphism proposed in this chapter is based on the orthogonality of the kernel matrix which makes it full rank. Furthermore, It is shown that the Q-rank of the quadratic form is bounded by the rank of the kernel matrix. This increases the difficulty of interpolating the coefficients of the core map.

5.2 Preliminaries

5.2.1 Core Map

The core map is a multivariate quadratic map defined mathematically as

$$\begin{aligned} f_1(x_1,\ldots,x_n) &= \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}^{(1)} x_i x_j + \sum_{i=1}^{n} f_i^{(1)} x_i + f_o^{(1)} \\ f_2(x_1,\ldots,x_n) &= \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}^{(2)} x_i x_j + \sum_{i=1}^{n} f_i^{(2)} x_i + f_o^{(2)} \\ &\vdots \\ f_m(x_1,\ldots,x_n) &= \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}^{(m)} x_i x_j + \sum_{i=1}^{n} f_i^{(m)} x_i + f_o^{(m)} \end{aligned} \tag{5.1}$$

They can also be defined as a function of the coefficients and univariate monomials as

$$\begin{aligned} f_1(x_1,\ldots,x_n) &= \sum_{0\le i,j<n} \alpha_{ij} X^{q^i+q^j} + \sum_{0\le i<n} \beta_i X^{q^i} + \gamma \\ &\vdots \\ f_m(x_1,\ldots,x_n) &= \sum_{0\le i,j<n} \alpha_{ij}^{(m)} X^{q^i+q^j} + \sum_{0\le i<n} \beta_i^{(m)} X^{q^i} + \gamma^{(m)} \end{aligned} \tag{5.2}$$

where $X^{q^i+q^j}$ are univariate monomials and X^{q^i} is the Frobenius automorphism over F_q.

5.2.2 Multivariate Quadratic problem

Definition 13 (MQ Problem)**.** *Given a polynomial* $f = f^{(1)}(x_1,\ldots,x_n),\ldots,f^{(m)}(x_1,\ldots,x_n) \in F_q[\,x_1,\ldots,x_n]$ *where* F_q *is a finite field of size* q *and* $F_q[\,x_1,\ldots,x_n]$ *is a polynomial ring with* m *equations of degree 2 and* n *variables, the goal is to find* $\bar{x} \in F_q^n$ *such that* $f(\bar{x}_o) = 0$.

The MQ problem has proven to be NP-Hard for quadratic polynomials over the field $GF(2)$[165]. If an invertible quadratic map is defined as $F\colon F^n \mapsto F^m$ where $F \in F_q^n$, then two invertible affine transformations $S\colon F^m \mapsto F^m$ and $T : F^n \mapsto F^n$ can be constructed to hide the algebraic structure of the map F. This map is a core component of the public key. By employing functional composition the public key is generated as $P_k = S \circ F \circ T$. Furthermore, constructing a canonical isomorphism of vector spaces as $\phi\colon F^n \mapsto F_q^n$, F is transformed into a quadratic map. Consequently,

this quadratic map is defined as $\tilde{F} = \phi^{-1} \circ F \circ \phi$. The result is a public key with the structure $P_k = S \circ \tilde{F} \circ T = S \circ \phi^{-1} \circ F \circ \phi \circ T \colon F^n \mapsto F^n$.

5.2.3 NTT

The Number Theoretic transform is a transform where the twiddle factors are sampled from a finite field. Point wise multiplication of a polynomial f and a polynomial g can be computed using number theoretic transform as

$$f.g = NTT^{-1}\omega(NTT\omega(f^{'}) \circ NTT\omega(g^{'})) \mod M \tag{5.3}$$

where ω is the primitive nth root of unity. Furthermore, the primitive root of unity can be written as $\omega^n = 1$ where $gcd(\omega, q) = gcd(N, q) = 1$ and M is the modulo of the polynomial. The modulo of the polynomial can assume any form. $f^{'}$ and $g^{'}$ are transformed from a degree n polynomials to $2n$ degree polynomials defined as $f = \sum_{i=0}^{n-1} \alpha_i X^i$ where $\alpha_i \in F_q$. For a transformed polynomial the invertibility probability is given as $1 - q^{-n}$[166]. The primitive root of unity can be expressed as $\omega = (a + ib)^{\frac{2^{m+1}}{N}}$ where $a \equiv 2^{2^{m-2}}, b \equiv -3^{2^{m-2}} mod(2^m - 1)$ [167]. Furthermore, NTT can be generalized to a set of integers of a finite field of q elements as follows[168];

Definition 14. *If the modulo is given as $q \equiv 1(mod 2n)$ were n is a power of 2 then F_q has 2n-th roots of unity ω^i where $i = 1, 3, \ldots, 2n - 1$ indexes the coefficients of polynomials $x - \omega^i \in F_q$.*

Example 1. *Let an irreducible polynomial be given as $1 + 3x^2 + 6x^4 + 3x^6 + x^8$, where the highest degree is $d = 8$. A prime given as $q = nk + 1$ where $k = 1$ is chosen. Zero padding the coefficients of the polynomial to length $q - 1$ results in $(1, 3, 6, 3, 1, 0, 0, 0)$. This is to make the reverse of the transform to be congruent to the original array of coefficient with a given choice of nth root of unity. Selecting the 8th root of unity in Z_9. Consequently, the forward NTT is given by Equation 5.11, 5.12 and 5.13.*

$$\begin{pmatrix} 2^0 & 2^0 & 2^0 & 2^0 & 2^0 & 2^0 & 2^0 & 2^0 \\ 2^0 & 2^1 & 2^2 & 2^3 & 2^4 & 2^5 & 2^6 & 2^7 \\ 2^0 & 2^2 & 2^4 & 2^6 & 2^8 & 2^{10} & 2^{12} & 2^{14} \\ 2^0 & 2^3 & 2^6 & 2^9 & 2^{12} & 2^{15} & 2^{18} & 2^{21} \\ 2^0 & 2^4 & 2^8 & 2^{12} & 2^{16} & 2^{20} & 2^{24} & 2^{28} \\ 2^0 & 2^5 & 2^{10} & 2^{15} & 2^{20} & 2^{25} & 2^{30} & 2^{35} \\ 2^0 & 2^6 & 2^{12} & 2^{18} & 2^{24} & 2^{30} & 2^{36} & 2^{42} \\ 2^0 & 2^7 & 2^{14} & 2^{21} & 2^{28} & 2^{35} & 2^{42} & 2^{49} \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 6 \\ 3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \equiv_9 \tag{5.4}$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 7 & 5 & 1 & 2 \\ 1 & 4 & 7 & 1 & 4 & 7 & 1 & 4 \\ 1 & 8 & 1 & 8 & 1 & 8 & 1 & 8 \\ 1 & 7 & 4 & 1 & 7 & 4 & 1 & 7 \\ 1 & 5 & 7 & 8 & 4 & 2 & 1 & 5 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 7 & 5 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 6 \\ 3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \tag{5.5}$$

$$\begin{pmatrix} 14 \\ 62 \\ 62 \\ 56 \\ 56 \\ 86 \\ 14 \\ 62 \end{pmatrix} \equiv_9 \begin{pmatrix} 5 \\ 8 \\ 8 \\ 2 \\ 2 \\ 5 \\ 5 \\ 8 \end{pmatrix} \tag{5.6}$$

To recover the array, the modular inverse is computed using Euler's theorem [169] where $a^{-1} \equiv a^{q-2} \mod q$*. The inverse NTT is applied with bit ordering. Removing the extra padded bits recovers the original vector.*

$$7^{-1}\begin{pmatrix}1 & 1 & 1 & 1 & 1 & 1 & 1 & 1\\ 1 & 2^{-1} & 4^{-1} & 8^{-1} & 7^{-1} & 5^{-1} & 1^{-1} & 2^{-1}\\ 1 & 4^{-1} & 8^{-1} & 1^{-1} & 4^{-1} & 7^{-1} & 1^{-1} & 4^{-1}\\ 1 & 8^{-1} & 1^{-1} & 8^{-1} & 1^{-1} & 8^{-1} & 1^{-1} & 8^{-1}\\ 1 & 8^{-1} & 4^{-1} & 1^{-1} & 7^{-1} & 4^{-1} & 1^{-1} & 7^{-1}\\ 1 & 5^{-1} & 7^{-1} & 8^{-1} & 4^{-1} & 2^{-1} & 1^{-1} & 5^{-1}\\ 1 & 1^{-1} & 1^{-1} & 1^{-1} & 1^{-1} & 1^{-1} & 1^{-1} & 1^{-1}\\ 1 & 2^{-1} & 4^{-1} & 8^{-1} & 7^{-1} & 5^{-1} & 1^{-1} & 2^{-1}\end{pmatrix} \tag{5.7}$$

$$\begin{pmatrix}5\\8\\8\\2\\2\\5\\5\\8\end{pmatrix} \equiv_9$$

$$\begin{pmatrix}1 & 1 & 1 & 1 & 1 & 1 & 1 & 1\\ 1 & 2 & 4 & 8 & 7 & 5 & 1 & 2\\ 1 & 4 & 7 & 1 & 4 & 7 & 1 & 4\\ 1 & 8 & 1 & 8 & 1 & 8 & 1 & 8\\ 1 & 7 & 4 & 1 & 7 & 4 & 1 & 7\\ 1 & 5 & 7 & 8 & 4 & 2 & 1 & 5\\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1\\ 1 & 2 & 4 & 8 & 7 & 5 & 1 & 2\end{pmatrix}\begin{pmatrix}5\\8\\8\\2\\2\\5\\5\\8\end{pmatrix} \equiv_9$$

$$7^{-1}\begin{pmatrix}7\\3\\4\\6\\1\\0\\7\\3\end{pmatrix} \equiv_9 \begin{pmatrix}1\\3\\1\\6\\7\\0\\1\\3\end{pmatrix} \equiv_9 \begin{pmatrix}1\\3\\6\\3\\1\\0\\7\\1\end{pmatrix}$$

5.3 Key recovery attacks

In this section, key recovery attacks employed to recover the affine isomorphism are reviewed and in the process a multivariate quadratic polynomial cryptosystem can be broken.

5.3.1 Linearization attack

Given a polynomial $f = (f_1, \ldots, f_m) \in F_q[\ x_1, \ldots, x_n]\ ^m$ that forms a linearization equation given as $\sum_{i=1}^{n} \sum^{m} \alpha_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + \sum_{i=1}^{m} c_i y_i + d \in F_q$. Given a multivariate polynomial based encryption algorithm, we have $\sum^{n} \sum^{n} \alpha_{ij} y_i y_j + \sum_{n} b_i^{(k)} y_i + \sum^{n-b} \sum^{n} \alpha_{ij}^{(k)} y_i y_j + c_i^{(k)} = 0$ where $\alpha_{ij}, b_i^{(k)}, \alpha_{ij}^{(k)}$ are unknown coefficients of the core map. The number of these coefficients are given in [170] as $n \sum_{j=0}^{n} \binom{m}{j} + m + 1 = n\binom{m+n}{n} + m + 1$ with a complexity of $O\left(n\binom{m+n}{n} + m + 1\right)^{\omega}$. To break an encryption scheme, the algebraic constant ω should be ≤ 2.8.

5.3.2 Minimum Rank Attack

The central polynomial $F = f^{(k)}(x_1, \ldots, x_n) \in F_q[\ x_1, \ldots, x_n]\ ^m$ can be exploited through Min. Rank attack with information on the rank of the linear combination of variables that make up F[171]. The goal is to find a solution $x \in F_q^m$ such that the rank of $F = \sum_{i=1}^{m} x_i f_i \leq 2$. In other words, $F = f^{(k)}(x_1, \ldots, x_{2n})$. The affine transformation S can be extracted if the number of equations m of F and P with low rank are computed. Extracting S can give the adversary the opportunity to retrieve the public key. If the variables that make up the two isomorphic maps T and S are small, an adversary can employ high rank attack to detect the last $n - b$ variables and the algebraic structure of the affine map S. For security against high rank attack , the number of queries of the last variable should be bounded by $q^n \geq 2^{80}$ for 2^{80} bit security[172]. The high rank attack sets the kernel of the central map to zero and checks for solution. Thereafter, it sets the transformed kernel for the coefficient that has dimension $n - 2n$. If the kernel has a probability that an inverse of the affine transformation T of a variable with rank $n - 2n$ gives the solution then T can be extracted, consequently the private key. The complexity of the Min. Rank attack on a core map is given as $O(q^{n.(r+v+a-1)}.(n-a)^3)$ where v is the number of vinegar variables.

5.4 New approach using NMNT

The New Mersenne number transform of the central polynomial $f(n)$ for a transform length $N = 2^k$ chosen for efficient random butterfly operation is given as

$$NTT(x) = \sum_{n=0}^{N-1} f(n)\omega_N^{\frac{n(2k+1)}{2}} mod(x^n + 1) \tag{5.8}$$

where $F(x) = \left[\sum_{n=0}^{N-1} \alpha_{ij}(\omega_N^q)^i\right]_{n=0}^{N-1}$. This is composed with the core map to generate the public key as follows $P_k = NTT(X) \circ F$. Such a multivariate cryptosystem can be described as an Isomorphism of polynomial with one secret[173]. The kernel matrix made up primitive root of unity and imbued with an orthogonal property is represented as follows

$$T_d = \begin{bmatrix} \omega_N^0 & \omega_N^0 & \omega_N^0 & \omega_N^0 \dots \omega_N^0 \\ \omega_N^{\frac{1}{2}} & \omega_N^{\frac{3}{2}} & \omega_N^{\frac{5}{2}} & \omega_N^{\frac{7}{2}} \dots \omega_N^{\frac{2N-1}{2}} \\ \omega_N & \omega_N^3 & \omega_N^5 & \omega_N^7 \dots \omega_N^{2N-1} \\ \omega_N^{\frac{3}{2}} & \omega_N^{\frac{9}{2}} & \omega_N^{\frac{15}{2}} & \omega_N^{\frac{21}{2}} \dots \omega_N^{\frac{3(2N-1)}{2}} \\ \vdots & \vdots & \vdots & \vdots \\ \omega_N^{\frac{N-1}{2}} & \omega_N^{\frac{3(N-1)}{2}} & \omega_N^{\frac{5(N-1)}{2}} & \omega_N^{\frac{7(N-1)}{2}} \dots \omega_N^{\frac{(2N-1)(N-1)}{2}} \end{bmatrix} \tag{5.9}$$

The rank of T_d is $N-1$ if and only if T_d has at least $N-1$ rows. To invert the transform, the transpose of the kernel matrix is multiplied by the inverse of the transform length to generate the original polynomial sequence given by

$$NTT^{-1}(x) = N^{-1} \sum_{n=0}^{N-1} F(x)\omega_N^{\frac{n(2k+1)}{2}} mod(x^n + 1) \tag{5.10}$$

The maximum degree of the polynomial sequence for transform length $N = 2^{k-1}$ is equivalent to $d = \frac{2^{k+1}}{k-1}$ and it is expected that the rank of the polynomial sequence should be bounded by the logarithm of the degree of the core map $log_q d$. However, this condition makes it easier to attack the core map. The Chinese remainder theorem describes the NMNT as natural ring isomorphism where the monomials $X^{q^i+q^j}$ are mapped into non linear orthogonal univariate polynomials $X^{\omega^i+\omega^j}$ over Z_q. In the nth iteration, $\sum_{i=0}^{n} \alpha_o \omega^{q^i+q^j}, \dots, \sum_{i_{n-1}=0}^{n-1} \alpha_{n-1}\omega^{q^{i_{n-1}}+q^{j_{n-1}}}$ is transformed into $\sum_{i=0}^{n} \alpha_o \omega^{q^i+q^j}, \dots, \sum_{i_{n-1}=0}^{n-1} \alpha_n \omega^{q^{i_n}+q^{j_n}}$ which results in a complexity of $N \log N$ for polynomial sequence of size $N+1$[174]. It can be inferred that, $f \mapsto f(\omega), f(\omega^3), \dots, f(\omega^{2n-1}))$ for coefficients $(\alpha_0, \dots, \alpha_{n-1})$. To find the interpolation of $f(n)$ at $\omega(X \equiv \omega \mod (X - \omega))$ the decomposition of the primitive root

of unity given in[167] is used as follows

$$f(x) \equiv b_k mod(X - a_k) : f(x) \equiv \sum_{k=0}^{n-1} \prod_{j \neq k} (X - a_j)) \tag{5.11}$$

$$\prod_{j \neq k} (a_k - a_j))^{-1} b_k mod \prod_{k=0}^{n-1} (X - a_k)$$

Furthermore, reducing $f(n)mod(X-\omega)$ is equivalent to evaluating $f(n)$ at $\omega(X \equiv \omega$ mod $(X - \omega))$. A key recovery attack on the orthogonality of the isomorphism can only be successful if the complement of the orthogonal kernel is projected to the codimension one subspace of the kernel with complexity $O(q^{log_q + a + 1})$ where a is a vinegar variable[175]. In masking the core map for security purpose, the Number theoretic transformed polynomials must be invertible. The necessary condition is that the inverse of the degree of the polynomial belongs to the ideals in the ring of the form $Z[x]/(f)$ where f is the degree n irreducible polynomial and the primitive root of unity ω satisfies the orthogonality condition[176] Ω.

$$\Omega_q = \sum_{i=0}^{n-1} \omega^{q^i} = n\delta(q)(\mathtt{mod} x^n + 1), q = 0, 1, \ldots, n-1 \tag{5.12}$$

In other words, if $q \neq 0$, then $(\omega^q - 1)$ must be a zero divisor of $\Omega_q(\omega)$. This means that the relation $(\omega^q - 1)\Omega_q(\omega) = 0$ where $(\omega^q - 1) \in Z[x]/(f)$. The implications of this is that if there was an actual zero divisor then the transform is non-invertible which leads to consequence in cryptanalysis. To reinforce its invertibility, then non of the representatives $r^{'}$ are zero divisors in $Z[x]/(f)$. This leads to the theorem which seeks to explain the impact of applying the full rank NMNT isomorphic map to the rank of the representative matrix of the quadratic form that make up the core map.

Theorem 10. *The Q-rank of the quadratic form is bounded by the rank of the orthogonal kernel T_d, if the dimspan$(F^{(0)}, \ldots, F^{(d-1)}) \geq d - 1$ where d is the degree of the core map.*

Proof. A matrix representation of the public key is defined as

$$\hat{F}_d = \begin{bmatrix} T_d & 0 \\ 0 & I_o \end{bmatrix} \tag{5.13}$$

The matrix representation of the core map is represented by the ith Frobenius power f^i. Let $F^{(0)}, \ldots F^{(d-1)} \in F^{(d+o)\times(d+o)}$ be the row vector of the matrix representation of the functional composition. If the row vector is multiplied by the orthogonal kernel matrix, it will produce the product of the matrix representation of the ith Frobenius power and the coordinate matrix representation of the

public key[177]. In other words, $(F^{(0)}, \ldots, F^{(d-1)})T_d = (\hat{F}_d f^i \hat{F}_d^T, \ldots, \hat{F}_d f^{d-1} \hat{F}_d^T)$. Recall that the functional polynomial composition of the public key is given as $P = NTT(x) \circ F$ which is equivalent to and as a function of the kernel $(P_0, \ldots, P_{n-1}) = (T_d F^{(o)} T_d^T, \ldots, T_d F^{(d-1)} T_d^T)$. Combining it with its matrix representation becomes $(P_0, \ldots, P_{m-1}\hat{F}_d) = (T_d \hat{F}_d f^i T_d^T \hat{F}_d^T, \ldots, T_d \hat{F}_d f^{d-1} T_d^T \hat{F}_d^T)$. This results to $\sum_{i=0}^{m-1} \hat{f}_d P_i = V f^i V^T$ where $V = T_d \hat{F}_d$ and $V f^i V^T$ is termed an identification of the form $X_i = \phi(\hat{x})^{q^i}$. □

Using an iterative approach in designing NTT algorithm, the coefficients of the polynomial are reversed using the function $Bitreverse(\alpha)$. This means that the structure of the coefficients are maintained while checking if the degree of the product of the monomial and the core map is bounded by an arbitrary degree using consistency check of Macaulay matrices. During the process of reducing the modulo, the Hensel remainder is congruent to $\alpha M \mod q$ where M is the modulo. In Montgomery reduction, input coefficients are greater than q^2 and less than qM. qM is a product of two unreduced words mapped to a residue which is less than $2q$ where $M > q$. In the process, the $\mod q$ operation is converted to $\mod M$ operation in which the scaling factor n^{-1} leads to efficient implementation up to $\frac{n}{2}$ modular multiplication. The standard representation can be produced by rounding up and multiplying with the prime number. The implementation can be done in constant time, If the representative less than $2q$ is subtracted from q. Thereafter, an arithmetic shift $N-1$ where N is the length of one word signed integer is then used to compute logical AND and the result is added to the representative[168]. If n is a coprime to the prime q and if two signed words x and y are transformed then there exist $x^{'} = xn^{-1} \mod q$, $y^{'} = yn^{-1} \mod q$. The high product becomes $x^{'}y^{'} = xyn^{-1} \mod q$. This is if and only if $x \equiv y(\mod q)$ where x is a representative of the class $y \mod q$. This translates to $\frac{x-y}{q} = E$ where E is a multiple of q[168],[178].

5.4.1 Modulo reduction

In reducing the modulo of the polynomial, a modulo $M = 2^N k + 1$ and a prime $q < M$ is selected. This leads to wide choices for word length where N is the transform length. The coefficient α is expressed as a product of two words $\alpha = x.y$ such that for word x,there is a range $0 \leq x \leq \frac{M}{2}$ and for word y, there is a range $0 \leq y \leq q$. Let a coefficient $\alpha \in F_q$, there exist a range $0 \leq \alpha < q\frac{M}{2} = k2^N + 1$ where k is the odd positive integer Proth number, N is

Algorithm 8 NTT

Input: $f \in R_p$, $\omega \in Z_p$
Output: $f^{'} \in R_p$ `in bit reversed order`
$m \leftarrow 0$
for $(i \leftarrow 0, i < \frac{m}{2}, i++)$ **do**
  $\omega \leftarrow (2.i)^{\frac{n}{m}}$
if m `is even` **then**
  for $(j = -m, j < m, j++)$ **do**
    $A_1 \leftarrow R_{educt} A[\ -j + \frac{m}{2^\alpha} + 1]\ \omega_{\frac{n}{2}}^2 \mod (x^{\frac{n}{2^\alpha}} + 1)$
    $B_1 \leftarrow R_{educt} B[\ l + j + \frac{m}{2^\alpha} + 1] \omega_{\frac{n}{2}}^2 \mod (x^{\frac{n}{2^\alpha}} + 1)$
    else
    $A_2 \leftarrow R_{educt} A[\ -j + \frac{m}{2^\alpha} - 1]\ \omega_{\frac{n}{2}}^2 \mod (x^{\frac{n}{2^\alpha}} + 1)$
    $B_2 \leftarrow R_{educt} B[\ l + j + \frac{m}{2^\alpha} - 1] \omega_{\frac{n}{2}}^2 \mod (x^{\frac{n}{2^\alpha}} + 1)$
    $f^{'}[\ -j + \frac{m}{2}]\ = A_1 - B_1$
    $f^{'}[\ l + j + m]\ = A_1 - B_1$
  `Return` $f^{'}$

Algorithm 9 INTT

Input: $f^{'} \in R_p$, $\omega^{-1} \in Z_p$, β^{-1}
Output: T_x
$m \leftarrow 1$
for $(i \leftarrow 0, i < \frac{m}{2}, i++)$ **do**
  $\omega \leftarrow (2.i)^{\frac{n}{m}}$
if m `is even` **then**
  for $(j = -m, j < m, j++)$ **do**
    $A_1 \leftarrow R_{educt} A[\ -j + \frac{m}{2^\alpha} + 1]\ \omega_{\frac{n}{2}}^2 \mod (x^{\frac{n}{2^\alpha}} + 1)$
    $B_1 \leftarrow R_{educt} B[\ l + j + \frac{m}{2^\alpha} + 1] \omega_{\frac{n}{2}}^2 \mod (x^{\frac{n}{2^\alpha}} + 1)$
    else
    $A_2 \leftarrow R_{educt} A[\ -j + \frac{m}{2^\alpha} - 1]\ \omega_{\frac{n}{2}}^2 \mod (x^{\frac{n}{2^\alpha}} + 1)$
    $B_2 \leftarrow R_{educt} B[\ l + j + \frac{m}{2^\alpha} - 1] \omega_{\frac{n}{2}}^2 \mod (x^{\frac{n}{2^\alpha}} + 1)$
    $T_x[\ -j + \frac{m}{2}]\ = A - B$
    $T_x[\ l + j + m]\ = (A - B)^{-2nl}$
for $l \leftarrow 0, l < i^2, l = i + n$ **do**
  $T_x[l]\ \leftarrow T_x[l]\ \beta^{-1} \mod q$
  `Return` T_x

the transform length and M is the modulo. If $f = dM + r^{'}q$ where $d \in F_q$ is the quotient, then the Hensel representative satisfies $-q < r^{'} < \frac{M}{2}$ for $0 \leq d < \frac{M}{2}q$. The goal is to reduce $fM \mod q$ and to compute $r^{'}$. This will eventually make the result congruent to $f \mod q$. First, it is pre-computed with $M^{-1} \mod q$ which gives $f = dM(\mod q)$. Consequently, $fM^{-1} = d(\mod q)$ becomes $d = fM(\mod q)$ and further computation results to $M^{-1}d = f(\mod q) = \beta m$. Also, there exist $k^2N = -1(\mod q)$, which implies that k^2N is prime and $-M^{-1}d - k^2N = 0(\mod q)$. This leads to $r^{'} = \frac{a-\beta m}{q} > \frac{Mq}{2M} = -q$. Assuming α is split into two words, then $\alpha = \alpha_o + \alpha_i M 2^N$ where $0 < \alpha_0 < M.2^N$ .Therefore, the representative becomes

$$r^{'} = \frac{\alpha - d}{qM} > \frac{\beta q k 2^N + 1}{2M.2^N} = -q \tag{5.14}$$

5.5 Experiment

5.5.1 Key generation

The NMNT was implemented in the Key generation and signature generation algorithm of the Round 1 Post Quantum submission for Gui algorithm[179] against the affine transform implemented in the algorithm using Array manipulation methods and bit-wise shift operations. The implementation was done in an AMD Ryzen 3 2200U with Radeon Vega Mobile Gfx graphic card with processor speed of 2500MHz, 2 cores, 4 logical processors and clock speed of 2.5GHz. The Operating system used was the Microsoft Windows 10 Home version 10.0.18362 with Build 18362 and there was no use of special instruction set. The results were recorded from the Windows Power shell. The variables of the core map were stored in blocks and copied to the temporary buffer of the output block before swapping the blocks in the reverse order. The transform size and the variables of the core map is declared as an unsigned integer type with 64bits size where the representative $r = 2^{64}$ for compiler optimization. The twiddle factors and the irreducible polynomials were pre-processed offline before feeding them into the algorithm. The parameters for implementing the signature generation scheme include; the degree extension m, the degree of the core map d, the number of equations a, the number of variables v and the repetition factor k. These parameters were adapted from [179]. The repetition factor k was chosen in order to reduce the rank of the linear combination of the variables that make up central polynomial. This would speed up the signature computation time. However, care was taken in order to make the rank as small to be easily decomposed by a Minimum rank attack. The hash function for

the signature generation algorithm was implemented using the Openssl Secure hash function library [180]. To avoid overflow during multiplication of the polynomials, the size of the array for storing the coefficients ranged from -9007199254740992 to 90071992547401894. From the results, it can be seen that the public key size using NMNT isomorphism is bigger than using affine isomorphism by a difference that increases with increasing field size. The increase in the public key size is due to the complexity of the butterfly operation in NMNT as a result of modular multiplication. This in turn increased the number of variables in the core map. In other words, NMNT modification of the core map comes at a cost of increased key size. There are different butterfly operations which can be applied and is effect on the key size noted. However, this is beyond the scope of the thesis. Furthermore, the verification time increased due to the complexity of field arithmetic as result of the number of variables in the core map and the selected prime. This can be reduced by using specialized processor instruction set. However, because of the absence of specialized instruction set in the implementation, we proposed reducing the complexity of butterfly operation by employing efficient modular reduction and multiplication methods like the Russian peasant multiplication [181]. In addition, the modular reduction process can be improved further and the twiddle factors and irreducible polynomials should be further reduced. Nevertheless, applying NMNT isomorphism performs better than affine isomorphism in terms of the secret key size and signature bits. The smaller key size is as a result of the sparse nature of the coefficient matrix of the core map. Finally, the public key size can be determined by the storage of equations and variables over the extension field. Generally, for reduction in the key size, the number of variables of the core map should be reduced as possible while keeping the degree constant. The timings of execution was also recorded and presented in Table 2. The time results is categorized by the platform of implementation.

Table 5.1: Key Sizes and Signature bits

Isomorphism	**m**	**d**	**a**	**v**	**k**	**pk(KB)**	**sk(KB)**	**signature bits**	
Affine	184	33	16	16	2	416.3	19.1	360	[43]
Affine	312	129	24	20	2	1955.1	59.3	504	[43]
Affine	448	513	32	28	2	5,789.2	155.9	664	[43]
NMNT	184	33	16	16	2	422.1	15.0	45	This work
NMNT	312	129	24	20	2	1990.0	41.8	63	This work
NMNT	448	513	32	28	2	5,903.4	94.7	83	This work

Table 5.2: Timings

Isomorphism	m	d	a	v	k	Key gen.	Sig.	Ver.	
Affine	184	33	16	16	2	343ms	16.1ms	0.057ms	[43]
Affine	312	129	24	20	2	2360ms	864ms	0.256ms	[43]
Affine	448	513	32	28	2	71,485ms	42,156ms	0.542ms	[43]
NMNT	184	33	16	16	2	43.72ms	20.0ms	58.889ms	This work
NMNT	312	129	24	20	2	61.02ms	45.2656ms	38.406ms	This work
NMNT	448	513	32	28	2	73.75ms	67.0491ms	31.484ms	This work

5.5.2 Minrank attack

To test the complexity of key recovery attacks on isomorphism, the Sidon cryptosystem was attacked and the implementation given in [44] was optimized for this purpose. The platform of implementation was the SageMathCloud [182] with 16GB of Random Access Memory, 3 shared CPU clusters and 20GB of hard disk. The mean success times of each computation was recorded for different prime values 2, 3 and 5 with degree $k \leq 12$. This implies that the number of variables $v = 21$ and number of equations $m = 24$. The plot is shown in Figure 5.1, 5.2 and 5.3 respectively. The execution of the attack was difficult to terminate at degree $k \geq 12$ after several hours. To compute the ideal of the linearized system using Gröbner basis, the F4 library [183] was used. From the results, it can be seen that at field with prime 2, the NMNT isomorphism performs better than the affine invertible isomorphism in terms of complexity to the Minrank attack with increase in the degree of the polynomial. However, on increase of the field size, both isomorphism have equivalent performance, for increasing degree. The presence of isomorphism with large kernel introduces some element of complexity when reducing the homogeneous equations to a linear equation where the ideal is generated through Gröbner basis operation. Homogeneous equations with degree 3 monomials are equivalent to multiples of leading monomials with non zero coefficients. A homogeneous polynomial $F(k, x_i)$ can be reduced to $x_{i+1} = \lambda x_i^d + \lambda_{n-1} x_i^{d-1} + \wedge(x)$ where $\wedge(x)$ is the vectorized representation of the isomorphism. If $\gcd(k, 2^n - 1) = 1$, then the non leading monomial of an element of the ideal becomes the permutation in the field F_{2^n} and also forms a basis to F_{q^n} over F_q. This also forms a linear combination with the central map. This permutation forms the number of solutions. The number of equations that lie in the plane of the isomorphism for $1 \leq i, j \leq k-2$ is given by $F(k, x_i) - y_i = k^2 - 2k - \binom{k}{2}$ where $\binom{k}{2}$ is the number of leading monomials of the form $x_i y_j$. If the coefficient of the leading term is not zero then $F(k, x_i)$ reduces to zero [44]. Further complexity is due to computing degree of regularity which is

explained in the next section. It can be said that for field size with prime $p = 2$, elements of the true solution with basis $(\beta_i\beta_j)$ for $i \leq j$ lies in the span of the kernel of the linearized equation for affine isomorphism than for NMNT isomorphism for increasing degree. This is shown in Figure 5.1

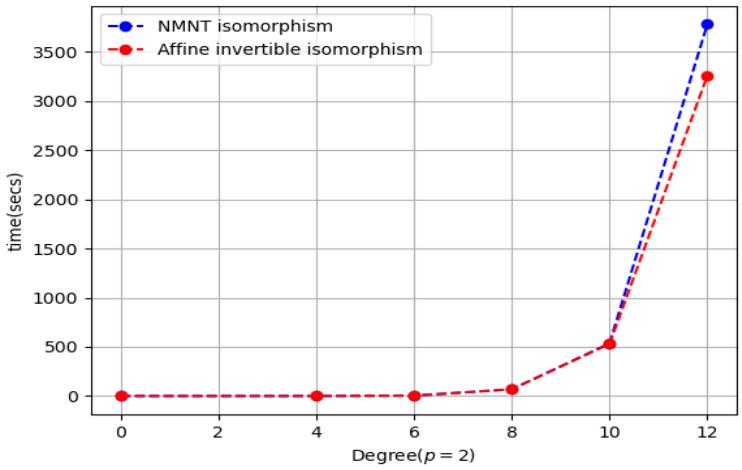

Figure 5.1: Complexity of Key recovery attack,field size $p = 2$

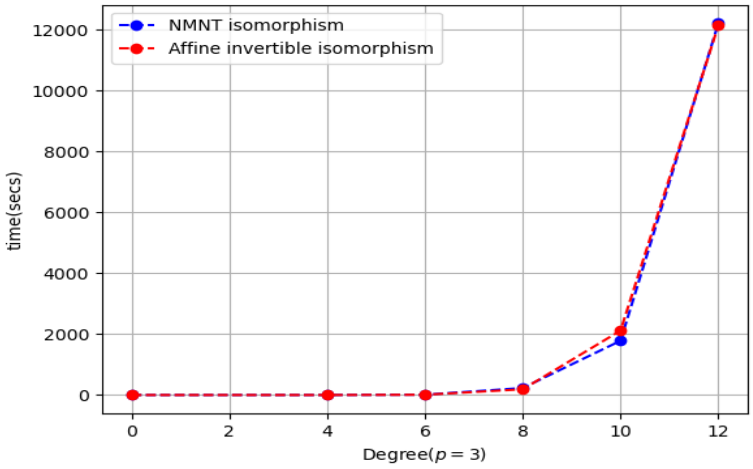

Figure 5.2: Complexity of Key recovery attack,field size $p = 3$

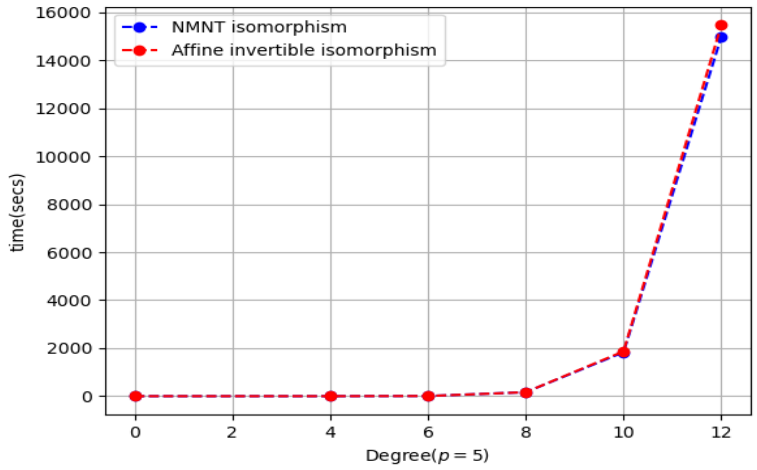

Figure 5.3: Complexity of Key recovery attack,field size $p = 5$

5.6 Multivariate Quadratic Solvability

In this section, we expound on the approaches of interpolating the core map of a quadratic polynomial cryptosystem. This approach works if the polynomial has been reduced to its ideal using an efficient reduction process [184].

5.6.0.1 Degree of Regularity

The definition of the maximal degree of m equations of a multivariate polynomial for an index $i \subset \{1, \dots\ m\}$ is given as

Definition 15. *For an index $i, 1 \leq i \leq m$, the maximal algebraic degree of the product of the coefficients of a map F: $F_2^n \mapsto F_2^m$ is given as*

$$\delta_k(F) = max_{i \subset \{1,\dots,m\}} deg(\prod_i F_i)$$

Given a homogeneous polynomial $h_1(x_1, \dots x_n), \dots h_m(x_1, \dots x_n)) \in F_q[\, x_1, \dots x_n]^{\,m}$ and a non-homogeneous polynomial $f_1(x_1, \dots x_n, \dots f_m(x_1, \dots x_n)) \in F_q[\, x_1, \dots x_n]^{\,m}$, a homogeneous ideal is defined as $I = \langle h_1, \dots, h_m \rangle$ and a non-homogeneous ideal is defined as $\tilde{I} = \langle f_1, \dots f_m \rangle$. The degree of regularity of a polynomial ideal I is defined as

$$\delta_i(F) = \{h_i \in F_q[\, x_1, \ldots x_n]\,^m / h \in I, deg(f) \leq d \binom{n+d-1}{d} \}$$

The degree of regularity of a polynomial ideal I in a semi regular sequence can be related to the first negative coefficient in the Hilbert series [185] H_s as

$$H_s = \sum_{i \geq 0} c_{ik} z^i = \frac{\prod_{i=1}^{m}(1 - z^{di})}{(1-z)^n} \tag{5.15}$$

when given degrees of regularity $d_i \forall 1 < i < m$ and where $\prod_{i=1}^{m}(1 - z^{di})$ is equivalent to the Hilbert function $HF_{m,n}$. For an invertible matrix $G_{(a)}$ over F_q, the distribution of the degree of variables in $G_{(a)}$ is given by

$$\lambda(n) = \prod_{i=1}^{m}\left(1 - \frac{1}{q^i}\right) \tag{5.16}$$

The expression that defines the relation between the dimension of the kernel and the differential of the function f with respect to the solution x is given in Lemma 17 [39]

Lemma 15. *The probability that the differential D_{xf} has kernel of dimension $i \geq 1$ given a function $f \colon (F_q)^n \mapsto (F_q)^k$ and where F_q^k is a k dimensional vector space with q^k elements and a solution $x \in F_q$ is given by*

$$\frac{\lambda(n)\lambda(n-1)}{\lambda(k)\lambda(k-1)\lambda(n-k)} q^{-k(k-1)} \tag{5.17}$$

A semi regular sequence is a sequence in which the homogeneous ideal does not belong to the finite field with q elements. For random undetermined systems, the sequence is regular. In multivariate cryptography, applying a generic change of coordinates to affine transformations that make up the private key or overdetermined systems with zero dimensions specifies the ideal. This is defined in Theorem 11 [185]

Theorem 11. *Let $I \subseteq F_q[\, x_1, \ldots, x_n]$ be a homogeneous ideal and $\tilde{I} \subseteq F_q[\, x_1, \ldots, x_n]$ a non homogeneous ideal with I defined by coordinates that are modified in generic terms in zero dimensions(overdetermined) then*

$$solv.deg_{DRL}(I) \leq \delta(\tilde{I}) \tag{5.18}$$

The degree of regularity of Gröbner basis is given by Theorem 12

Theorem 12. *Let d_{reg} be the degree of regularity of the polynomial F of ZHFE(q, n, D). If $q > 2, D \leq q + 2qs$, then*

$$d_{reg} \leq \frac{(q-2)(s+2)}{2} + 2 \leq \frac{(q-1)\lfloor log_2 D \rfloor + 3}{2} + 2 \tag{5.19}$$

If $q = 2, D \leq 2 + 2^{s-1} + 2^s$ then

$$d_{reg} \leq \frac{s+3}{2} + 2 \leq \frac{\lfloor log_2 D \rfloor + 4}{2} + 2 = \frac{1}{2}\lfloor log_2 D \rfloor + 4. \tag{5.20}$$

Theorem 13 improves on the bound of the algebraic degree from the result in [186] which states that for F a permutation over F_2^n, there is an integer k and l such that $\delta_k(F) < n - l$. Review of the algorithm is given as follows; Let $h\colon \mapsto \prod_{i \in k} F_i(x)$ with $|k| \leq k$ then for any $l \subset \{1, \ldots, k\}$, show that $\alpha_{ijk} = 0$ for a monomial $\prod_{j \notin l} x_j$. It follows that $\alpha ijk = \sum h(x)$ where $x \in F_2^n$ is such that $x_j = 0$ for $j \in l$ and the permutation $F(x) = 1 \mod 2$ for $i \in k$. Since $y = F(x)$ where $y \in F_2^n$ then $Fj^{-1}(y) = 0$ for $j \in l$. Consequently, a function $\pi_{k,l}$ is defined as $\pi_{k,l}\colon \begin{cases} x_{i \in k} \mapsto F_2 \\ x \mapsto \prod_{i \in l}(F_i^{-1}(x)) \end{cases}$ where $\pi_{k,l}$ is a function of $n-l$ variables and $\delta_k(F) < n-l$.

Theorem 13. *Given a function $h\colon F_q^n \mapsto F_i(x)$ A multivariate system with m equation and n variables can be solved with degree bounded by $degF_i(x) \leq n - k$, if $\delta_k(F) \geq \frac{n-1}{n-k}$ $\forall monomial x_j$.*

Proof. Let $x \in F_2^n$ such that $x_j = 0, F_j(x) = 0$ and for $0 \leq i, j \leq k$. Let a function $h\colon \begin{cases} F_q^n \mapsto F_i(x) \\ F_q^n \mapsto \prod_n^k (f_i(x_i), \ldots, f_m(x_n)) \end{cases}$ where $F_i(x) = \sum_{i,j \leq k}^n \alpha ijk x_j x_k + \sum_j^n \beta_{ij} x_j + \gamma_i$ for coefficients $\alpha_{ijk}, \beta_{ij}, \gamma_i \in F_q$ and degree $degF_i(x) \leq n - k$. From theoretical results in [186],

$$F_i(x) = \sum_{x \in F_q^n} (-1)^{\beta_{ij} F(x) + \alpha_{ijk}.x} \tag{5.21}$$

This defines the Walsh coefficient and dividing the equation with 2^k gives

$$F_i(x) = \sum_{x \in F_q^n} (-1)^{\beta_{ij} F(x) \left(\mod 2^{\frac{\lceil n-1 \rceil}{degF(x)}+1} \right)} \tag{5.22}$$

It follows that a linear combination of the quadratic form becomes

$$\alpha_{ijk}.x: \begin{cases} \prod_j^n F_i(x) \mapsto & \mod 2\frac{\lceil n-1 \rceil}{degF(x)} + 1 \forall x_j = 0 \\ & x \mapsto F_j(x) \end{cases} \tag{5.23}$$

This results to

$$\mod 2\frac{\lceil n-1 \rceil}{degF(x)} + 1 = 0 \tag{5.24}$$

$$\frac{n-1}{n-k} \leq 1 \mod 2 \tag{5.25}$$

□

5.6.0.2 Evaluation of f & F

Let a set of multivariate polynomials be defined as $f^{(k)}(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \alpha_{ij}^{(k)} x_i x_j + \sum_{i=1}^{n} \beta_i^{(k)} x_i + c_i^{(k)}$ where $F = (f_1, \ldots, f_m) \in F_q[\ x_1, \ldots, x_n]^{\ m}$ and the coefficients $\alpha_{ij}^{(k)} x_i x_j, \beta_i^{(k)} x_i, c_i^{(k)} \in F_q$. The function f with variables $y_1 = y_{n-b+1}, \ldots, y_n$ containing the last b variables where $f^{(k)}(y_{n-b+1}, \ldots, y_n) = \sum_{i=n-b+1}^{n} \sum_{j=i+1}^{n} \alpha_{ij}^{(k)} y_i y_j + \sum_{i=n-b+1}^{n} \beta_i^{(k)} y_i + c_i^{(k)}$ where $f^{(k)}(y_{n-b+1}, \ldots, y_n) \in F_q^n$ can be evaluated which gives an isomorphic map . The function F again with $y_2 = y_1, \ldots, y_n$ containing the first $n-b$ variables and the last b variables can be evaluated. It follows that $f^{(k)}(y_1, \ldots, y_n) = \sum_{i=1}^{n-b} \sum_{j=n-b+1}^{n} \alpha_{ij}^{(k)} y_i y_j$. Transforming the function F with y_i now becomes $\tilde{F} = (f_1(y_{n-b+1}, \ldots, y_n), \ldots f_m(y_{n-b+1}, \ldots, y_n)) \in F_q[\ x_1, \ldots, x_n]$. This results to,

$$\tilde{F} = (f_1(x_1, \ldots x_{n-b}, y_1, \ldots y_n), \ldots f_m(x_1, \ldots f_m(x_1, \ldots, x_{n-b}, y_1, \ldots y_n) \in F_q[\ x_1, \ldots \tag{5.26}$$

$$x_{n-b}]^{\ m}$$

which satisfies the equation

$$\sum^{n} \sum n \alpha_{ij}^{(k)} y_i y_j + \sum^{n} \beta_i^{(k)} y_i + \sum^{n-b} \sum^{n} \alpha_{ij}^{(k)} y_i y_j + c_i^{(k)} = 0 \tag{5.27}$$

where $1 \leq i, j \leq \frac{nb(b+1)}{2}$. Let the coefficients $\alpha_{ij}^{(k)}, \beta_i^{(k)}, c_i^{(k)}$ be the linear combination equations of the columns of the Macaulay matrices A, the condition is checked whether $uA = (F \circ \tilde{F}) = 1$

5.6.0.3 Complexity

The F5 algorithm is employed to cryptanalyze multivariate cryptosystem after reducing the non linear system to a linear determined system through extraction of variables in as many iterations as possible. The complexity of solving the

system of $n(n-b)$ equations $n-b+1$ variables over F_q is given as $q^{n-b}O((n(n-b)\binom{n+dreg-1}{dreg})\Big)^{\omega}$ where ω is the measure of entropy with $2 \leq \omega \leq 3$. The linear algebra constant and q^{n-b} is the probability of guessing a solution to $n-b+1$ variables. If the number of equations is equivalent to the number of variables for any positive ϵ, in a semiregular system, a deterministic approach to testing consistency of the Macaulay matrices and employing the quantum oracle gives a bound on the complexity as $O(2^{(0.47+\epsilon)n})$ quantum gates while a probabilistic approach gives a bound as $O(2^{(0.462+\epsilon)})n$ quantum gates [38]. By extension, based on the complexity for an r string, a semiregular system is given by $O(2^{(1-2r+2F\alpha(r)\epsilon)n})$ for a deterministic approach. The probabilistic approach is given by $O(2^{(\frac{1-r}{2}+2F\alpha(r)+\epsilon)n})$ where $\gamma = 1-\frac{k}{n}, F\alpha(r) = r\log(D^d(1-D)^{(1-D)})$ with $D = M(\frac{\alpha}{r})$ and $M(x) = -x+\frac{1}{2}+\frac{1}{2}\sqrt{2x^2-10x-1+2(x+2)\sqrt{x(x+2)}}$. Therefore, the complexity becomes

$$O(q^{n-b}(n(n-b)n^{\omega}+2^{0.47+\epsilon)n}) = O(q^{n-b}(n(n-b)^{(0.4+\epsilon)\omega}) = O(q^{n-b}m^{(0.4+\epsilon)\omega}) \tag{5.28}$$

The asymptotic cost component of employing brute force search to guess a variable when a function belonging to a quadratic map is given by $log_2 q$ while that of the Grover's algorithm is given by $0.5log_2 q$ which is half of the cost exponent obtained when using brute force search. In the probabilistic approach, the cost component is given by $2^{(0.4+\epsilon)n}$ as n towards infinity, therefore the complexity becomes $O(q^{n-b}m0.5(log_2 q)^{\omega})$.

5.6.1 Sparse Approximation Solution

In this section, a sparse approximate approach to solve the reduced basis using the gradient descent method is presented. The major operation is to compute $f^m(\bar{x})$, where f^m is the sparse matrix which is a row echelon form of the matrix representing the non-linear system. If the number of non-zero elements is given by $[f^m(\bar{x})]_{>0}$, then the complexity of solving the sparse system using the gradient method becomes $O([f^m(\bar{x})]_{>0})$.

5.6.1.1 Problem Formulation

The optimization problem is formulated as

$$||f^m(\bar{x})||_F = \frac{1}{2}x^T x + \partial f(x)^T x + ||u(x)v(x)||_F^2 = \frac{1}{2}x^T x + \tag{5.29}$$

$$\partial f(x)^T x + \frac{1}{N}\sum_{i=1}^{N} u_i(x)v_i(x) - 2\sum u_i(x)v_i(x)$$

$$= \frac{1}{2}x^T x + \mathrm{Tr}(D(x)^T x D(x)) + \frac{1}{N}\sum_{i=1}^{N} u_i(x)v_i(x) -$$

$$2\sum_{i=1}^{N} u_i(x)v_i(x)$$

where $D(x)$ is convex and differentiable which signifies that $\forall x, y \in X, \forall t \in [\,0, 1]\; P(tx+ (1-t)y \leq tf(x) + (1-t)f(y)$. If $P(x)$ is differentiable, it signifies that $\partial F(x)$ has a definite gradient at x [187]. The term $u(x)v(x)$ defines the gradient vector.

5.6.1.2 Convergence

The Wolfe-Powell's rule [188] is employed to compute the descent and test for convergence. The descent direction d_k can be obtained for the objective function $||f^{(m)}(\bar{x})||_F$ which is a trace of the multivariate polynomials that contain the root of the equation. This is done by finding the negative definiteness of the differential of the core map $F(x)$ which is $-\partial F(x)$. In order to converge to a solution, the optimization problem formulated would need to be reduced accordingly by employing a factor (step length) which describes the difference between the original and reduced objective function. The optimization problem is further mapped to a tangent line at the root $\bar{x}$, where the tangent line points in the direction of the decent d_k. A choice of step length is paramount in order to ensure convergence. The reduced objective function is bounded by

$$||f^{(m)}(\bar{x}) + \alpha d_k)||_F \leq ||f^{(m)}||_F + \beta\alpha\partial f^{(m)}(\bar{x})^T d_k \tag{5.30}$$

where α is the step length and $0 < \beta < 1$. A check for optimal solution is carried out as a consequence of the updated optimization problem as the step length increases.

$$f^{(m)}(\bar{x}+\alpha d_k) = f^{(m)}(\bar{x}) + \beta\alpha\partial f^{(m)}(\bar{x})^T d_k \tag{5.31}$$

$$d_k = f^{(m)}(\bar{x}+\alpha d_k) - (f^m(\bar{x}) + \beta\alpha\partial f^{(m)}(\bar{x})^T) \tag{5.32}$$

$$\frac{\partial d_k}{\partial f^m} = \partial f^m(\bar{x}+\alpha d_k) - \partial f^m(\bar{x}) + \beta\alpha\partial^2 f^{(m)}(\bar{x})^T \tag{5.33}$$

$$= \partial f^m(\bar{x}+\alpha d_k) - \partial f^{(m)}(\bar{x}) + \beta\alpha \mathtt{H}^T \tag{5.34}$$

$$= \partial f^{(m)}(\bar{x}) + \partial f^m(\alpha d_k) - \alpha f^m(\bar{x}) + \beta\alpha H^T \tag{5.35}$$

$$= \partial f^m(\alpha d_k) + \beta\alpha H^T \tag{5.36}$$

$$= \alpha(\partial f^m d_k + \beta H^T) \tag{5.37}$$

For optimal solution $\frac{\partial d_k}{\partial f^m} = 0$. This results in

$$\alpha(\partial f^m d_k) = \alpha\beta H^T \tag{5.38}$$

where H is the hessian matrix. This verifies the assertion that if the step length α is controlled and given suitable choice of $\beta \in \{0, 1\}$, an optimal descent can be computed. This improves the distance of deviation of the Hessian matrix from $\mathrm{O}(p^3)$ flops to $\mathrm{O}(p)$. The condition of optimality can be extended to the Mercer's theorem [189] in which $\alpha\&\beta$ will solve the linear system

$$\left[\begin{array}{c|c} 0 & 1^T \\ \hline \vec{1} & \Psi + \tau^{-1} I \end{array}\right] \left[\begin{array}{c} \beta \\ \hline \alpha \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline y \end{array}\right] \tag{5.39}$$

where $(y_1, \alpha_1), \ldots (y_N, \alpha_N)$ and α_i is the step length. Since the core map is differentiable, it can be written that

$$\partial f^m(\bar{x}+\alpha d_k) - \partial f^m(\bar{x}) + \beta\alpha\partial^2 f^{(m)}(\bar{x})^T \leq \delta||(\bar{x}) - (\bar{x}+\alpha d_k)||_2^2 \tag{5.40}$$

Consequently,

$$\partial f^m(\bar{x}+\alpha d_k) \geq \partial f^m(\bar{x}) + \beta\alpha\partial^2 f^{(x)}(\bar{x})^T(\bar{x}+\alpha d_k) - (\bar{x}) + \frac{\delta}{2}||(\bar{x}+\alpha d_k) - (\bar{x})||_2^2 \tag{5.41}$$

The choice of the 2-norm is to improve convergence in the presence of variables in the input space of a set of solutions. If the hessian matrix H was to deviate by a distance $||\Lambda||_F^2$, then the variable δ would satisfy $\delta \leq \frac{||\Lambda||_F^2}{(NM)^2}$. It is recommended that the standard deviation in relation to the distance $||\Lambda||_F^2$ be kept minimum. This is analogous to the proposal by [190]. The update of the descent d_k as an expression of the Hessian matrix can be computed as follows.

$$\frac{1}{2}(H_{ij}^2 + H_{ii}H_{jj}\varphi^2 + (x_{ij} - H_{ij} + h_i^T D h_j)\varphi + \lambda|\bar{x_{ij}} + d_k + \varphi| \tag{5.42}$$

which results to

$$Tr(Hd_k^{'}Hd_k^{'}) = Tr(Hd_k W d_k) + 2\varphi_i w_i^T d_k w_i + \varphi^2(H_{ii})^2 \tag{5.43}$$

.

The optimization problem is therefore reduced as a result of the step length as follows

$$||x^T \partial F x||(1 - \frac{\alpha}{\beta} - ||uv^T||(\frac{\alpha}{1+\beta}) \tag{5.44}$$

$$\leq ||x^{\partial} F x||(1 - \frac{\alpha}{\beta}) - ||u||(\frac{\alpha}{1+\beta}.||v||(\frac{\alpha}{1+\beta}) \tag{5.45}$$

$$\leq ||s^T v||_F^2(\frac{\beta+\alpha}{\beta}) - ||\partial F^T v||_F^2(\frac{\beta-\alpha}{\beta}) \tag{5.46}$$

$$\leq ||xv||_F^2 \epsilon_o(\alpha+\beta) - ||\partial F^T v||_F^2 \tag{5.47}$$

$$\leq \epsilon_o(\beta^2 - \alpha^2)(||xv||_F^2 - ||\partial F^T v||_F^2) \tag{5.48}$$

$$\leq \epsilon_o(\alpha+\beta)(\alpha-\beta)(||xv||_F^2 + \sum_{i,j} \partial F^T ||v|| - \tag{5.49}$$

$$\frac{\Xi}{2}\sum_{i,j} ||xv||_F^2 ||\partial F^T v||$$

$$\leq \epsilon_o(\alpha+\beta)(\alpha-\beta)(||xv||_F^2 + \sum_{i,j} \partial F^T ||V|| - \frac{\Xi}{2} Tr(x^T \partial F v) \tag{5.50}$$

5.6.1.3 Minimizer

The minimizer defines a point where the solution exists. It can be constructed from $Tr(S^T \partial P V)$ and $Tr(\partial P^T V)$. Let variables $A = \frac{1}{2}(S + V^T)$ and $B = \frac{1}{2}(\partial F + V^T)$ be defined and a Jacobian matrix $J = B^{-1}A$. Consequently, A and B is transformed to new variables a and b as follows

$$a = Tr(\bar{x}^T A \bar{x})b = Tr(\bar{x}^T B \bar{x}) \tag{5.51}$$

The new variables are differentiated as follows

$$\begin{aligned} da &= 2|A\bar{x}||d\bar{x}| \\ db &= 2|B\bar{x}||d\bar{x}| \end{aligned} \tag{5.52}$$

By constructing a function $\Xi = ab$, further differentiation results in

$$d\Xi = b(da - \Xi db) = 2b|A\bar{x}||d\bar{x}| - 2\Xi|B\bar{x}||d\bar{x}| \tag{5.53}$$

Consequently, the descent becomes

$$\frac{\partial \Xi}{\partial \bar{x}} = 2b|A\bar{x}| - 2\Xi|B\bar{x}| \tag{5.54}$$

On conditions of optimality

$$\frac{\partial \Xi}{\partial \bar{x}} = 0 \tag{5.55}$$

it follows that

$$2b||A\bar{x}|| = 2\Xi||B\bar{x}|| \tag{5.56}$$

$$bJ||\bar{x}|| = \Xi||\bar{x}|| \tag{5.57}$$

$$bJ\lambda\vec{1} = \Xi\lambda\vec{1} \tag{5.58}$$

where λ is the eigenvector and $\vec{1} \in [\ 1, \dots 1^N]$. Equation (5.66) must satisfy the following conditions

$$bJ\begin{cases} = \lambda_{ij} & \text{if} \quad \bar{x}_{ij} > 0 \\ \in [\ -\lambda_{ij}, \lambda_{ij}] & \text{if} \quad \bar{x} < 0 \\ = -\lambda_{ij} & \text{if} \quad \bar{x}_{ij} = 0 \end{cases} \tag{5.59}$$

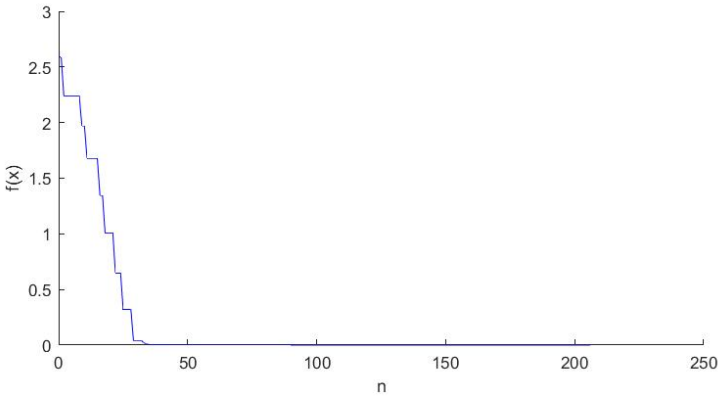

Figure 5.4: Quadratic map function $f(x)$ in (32) and (33) reduced to $1.3614e-12$ with gradient method after reduction to its ideal after iterations $n = 365$

Chapter Summary

In this chapter, the New Mersenne Number transform was used as a masking function in the key generating process of multivariate based cryptosystem to hide the structure of the quadratic forms that make up the core map. The kernel function of the New Mersenne Number transform can be functionally composed with the central core map to generate the public key used for encryption in a multivariate polynomial cryptosystem. This can be termed as a cryptosystem with one secret. The kernel function is orthogonal and can be inverted, if the inverse of the transform length is multiplied with it. The rank of the kernel is a function of the degree of the core polynomial sequence. However, because of this bound, it can be easily inferred, if the linear combination of the kernel is equal to its rank. This is a necessary condition for its attack by key recovery algorithms. Furthermore, another condition worthy of note is the image of the codimension one subspace of the kernel. The inverse of the degree of the core polynomial sequence belongs to the ideals in the ring. This would enable the kernel of the New Mersenne Number transform to be invertible. In addition, the residue must not be zero divisors in the ring. An isomorphism which is robust against key recovery attacks was generated. This isomorphism is complex against permutation and inversion as compared to affine isomorphism. It is also robust against key recovery attacks by reason of its orthogonal property which makes it full rank. A modified version of Montgomery approach was applied to reduce the modulo of the form $M = 2^n k + 1$. From the result of applying the proposed method as a map to the Gui Post quantum algorithm, a considerable reduction in secret key size and signature bits was observed. Furthermore, the complexity of using the key recovery attacks on isomorphism was presented using an implementation of the Sidon cryptosystem. It was discovered that both transforms realized equivalent performances against the attack with increasing degree as the field size scales up. This is due to the nature of the homogeneous non-linear equations whose ideal has been reduced using variant of Gröbner basis. Also, the degree of regularity parameter also plays a significant role in this observation. This parameter was analyzed and discussed in detail in this chapter. Finally, a sparse approximate solution of the ideal using gradient descent method was proposed. If the number of non-zero elements is given by $[f^m(\bar{x})]_{>0}$, then the complexity of solving the sparse system using the gradient method is given as $O([f^m(\bar{x})]_{>0})$.

Chapter 6

Conclusion and Further work

6.1 Conclusion

The sigmoid function of support vector machine kernel is chosen which defines the non-linear separability of the points in Z^n. The columns of an arbitrary chosen noise matrix was employed to expand the log likelihood ratio. This constructs a function that maps the vectors from the linear subspace to the low dimensionality linearly dependent subspace. Approximating this function would result to $O(n^2)$ operations. The lattice basis vectors converges if the projection and the affine transformation concatenate to find a solution to the optimization problem. Consequently, the basis vectors are projected into linear dependent positions, to meet the condition $\lim_{i\to\infty} ||\lambda_{k,i} - x_{k,i}|| = 0 \forall k, i \in F_q$. Bit Error rate results of using the proposed method was obtained. The results were compared with variants of the LLL reduction method. The lattice basis in this experiment was a Quadrature Amplitude Modulation channel matrix. From the results, the proposed approach gave a 1db improvement over the LLL algorithm. It also outperformed the Orthogonalization and QR lattice reduction methods. Furthermore, It was noted that the statistical distance between the lattice point distribution and the ideal Gaussian distribution should be as minimum as possible in order to analyze the scheme's security under the indistinguishability criterion. This can be achieved, if the target vector is sampled from a distribution $z_i \leftarrow D_{Z,\sqrt{\Sigma_i},t_i}$, spaced at the norm r_i and $|c_i - t_i|$. $|c_i - t_i|$ is the distance of c_i from a plane in the distribution. In addition, a new basis vector can be sampled, if U a random number generated from a uniform distribution $[\ 0, 1)$ satisfies the condition $U > \oslash[\ t]$ where $\oslash[\ t]$ is the cumulative distribution function and t is the result of reduction of the basis $\tilde{b}_i$ with a norm

$\|\tilde{b}_i\|$. Furthermore, to ensure λ-bits of security, the floating point precision should be bounded by $m^{'} \geq \lambda + log_2 \sum g_{i,i}(\mu^2 \tau sqn^{3/2})$ where μ^2 is the precision of the perfect sampler and n is the number of uniform vector samples. Finally, the precision is inversely proportional to the success probability of withstanding an attack and directly proportional to the security parameter.

It was observed that by employing Gaussian elimination and taking an instance of the boundary map, the generator matrix that represents the isomorphism can be reduced into its row echelon form. Furthermore, this reduction is a function of the ordering of the Plücker coordinates. The enumeration over the basis terminates, if the codewords with minimum distance $d_{min} \leq q^{\frac{k(k-1)}{2}}$ is found. Given the size of the coordinate, there is a bound on the enumerator with respect to the number of vectors that are indexed in a linearly independent set. In order words, using the analogy of section 4.3, there is a bound on the path weight of the subgraph induced by the intersection of the subspaces. Subsequently, the probability that k vertices with rank r has $q^{i(i-1)/2}$ adjacent vertices is given by $b_{r_k} \leq q^{i(i-1)/2} \begin{bmatrix} n \\ k \end{bmatrix}_q$. This was proved in Theorem 5. This probability depends on the set that contains $m - i$ elements of the syndrome space. Theorem 8 was employed to derive the bound on the number of monomials to be evaluated. The complexity of solving the syndrome decoding problem in the Grassmann metric was given by $O(\frac{(n-k)^2}{2} q^{\frac{k(k-r)}{2}(n-k)})$. From the results on the experiment on probability of failure, the non-negative Grassmannian code performed better than the LDPC code under IND-CPA model. This is because of the random choice of Plücker coordinates and the better pruning of the generator matrix that represents the isomorphism. Furthermore, on cost of Enumeration, Non-negative Grassmannian performed better than the LDPC code for smaller field sizes. The Plücker coordinates of the codewords of the parity matrix of the Non-negative Grassmannian code is assumed to be linearly independent and the vector with the minimum distance are also assumed to be sampled from a Gaussian distribution. The dependence between subspaces is modelled using copula functions. The function can be expressed as the degree distribution of the boundary measurement map as follows.

$$f(x/G_i) = \prod_{i=1}^{n} f_i(x_i) \times \prod_{i,j \in E} C_{i,j}(U_i, U_j) \tag{6.1}$$

From Lemma 12 in chapter four, it was shown that if $x \not\subset \texttt{supp}(C)$, then $\lambda(P_{V'}(F_K(v) \cap \texttt{supp}(C)) = 0$. From the results, the number of errors is directly proportional to the p columns of the Non negative Grassmann code. Also, the lower the bit error rate,

the higher the probability of solving the syndrome decoding problem. The shortest vector problem can be generalized to the syndrome decoding problem. To find the syndrome using its probability distribution, the distribution of the error coordinate positions is modeled using copula functions. For a copula function, if there exists index k such that the component of the transformation sampled from a uniform distribution at the kth position is zero, then the copula function becomes zero.

It was proven that the q-rank of the quadratic form of the core map is bounded by the rank of the kernel function, if the dimension of the map is greater or equal to its degree. Using solution check of the Macaulay matrices, the bound on the product of the degree of the monomial and the core map can be derived. A variant of Montgomery reduction was employed to reduce the modulus of the polynomial sequence. It was also noted that this transform works for monic polynomial sequence with rational coefficients. The coefficients is expected to have a size that is greater than the square of the field size and less than the product of two unreduced words. The selected modulo was $M = 2^N k + 1$ where N is the transform length. In reducing the modulus, the rational coefficients are expressed as the product of two words with varied ranges. The reduction process reduces $fM \mod q$ and computes a Montgomery residue. Experiments comparing the affine isomorphism with the kernel of New Mersenne Number transform was conducted using the Key generation algorithm of a variant of the Hidden Field scheme. From the result, NMNT performed less than the affine transform in terms of the cost of the public key which is due to the complexity of the butterfly operation. However, it performs better in terms of the secret key cost and signature bits. This is because of the size of the kernel of the NMNT as compared to the size of the affine matrix. The size of the kernel can be reduced further by manipulating the primitive roots of unity. Furthermore, the complexity of the Min-rank attack using the kernel of the NMNT isomorphism was compared to that of the affine isomorphism. Sidon spaces based cryptosystem was used in this experiment. The ideal of the linear system was computed using Gröbner basis. From the results, it can be seen at small field sizes, the kernel of the NMNT isomorphism performs better than the affine isomorphism. This is because of the size of the kernel of the affine isomorphism which scales poorly when system of non-linear equations is transformed to a linear system. Also, the computation of the degree of regularity can impact on its complexity to the Min-rank attack. These non-linear equations have multiples of leading monomials with non-zero coefficients. A basis in the field can be formed if $gcd(k, 2^N - 1) = 1$ where k is an index. The

relationship between the degree of regularity and the first negative coefficient in the Hilbert series was defined. The last b variables of the core map was evaluated which generates an isomorphism. If the number of equations in the linear system is equal to the number of variables, there are different bounds on the quantum oracle which depends on the approach. In other words, whether the approach is deterministic or probabilistic. The cost of guessing a solution to the $n-b+1$ variable using Grover's search is twice that of guessing using brute force method. This leads to complexity in computing the Gröbner basis using the F5 algorithm. To solve the sparse reduced basis, a gradient descent method was employed. The gradient descent method used Wolfe-Powell's rule to test for convergence. The complexity of using this iterative method was obtained as $O([f^m(\bar{x})]_{>0})$.

6.2 Future work

In this thesis, the shortest vector oracle has been the Gaussian sampler and its precision has been analyzed using the convolution theorem [26]. Further research effort will concentrate on applying the dimensionality mapping lattice reduction algorithm to other oracles such as enumeration with pruning [191] over mapping sets given random basis. Most importantly, the dimensionality mapping approach using linear programming approach whereby a mapping function is reduced to solving an optimization problem might not be the best approach as was observed from the result of the BER experiment. This is because such equations contains several variables and unknowns which affects computation time. In the future, the proposed method would be modified so as to eliminate the solution to an optimization problem step completely.

Furthermore, parameters for the basis norm where derived in Equations 3.10, 3.11, 3.12 and 3.17. Further derivation of the norm will take into consideration how the Hermite constant γ and the lattice volume scales with each block of the lattice vector. Future experimentation will take into consideration the orthogonality factor of the basis as a criteria of the quality of basis which is very important for cryptography. This is because the BER experiment basically tests the capacity of the basis to be used in solution to signal processing problems. In addition, other lattice reduction algorithms that are most especially applicable to cryptography would be considered most especially the Blockwise Korkine-Zolotarev (BKZ) reduction method [192]. The BKZ approach uses a block-size parameter that scales with reduction in lattice dimension or the rank of the lattice basis to produce clean basis

with better approximation factor. Furthermore, the convolution theorem needs some improvement to produce significant results on statistical distance.

In the solution to the syndrome decoding problem, probabilities of enumerating the bases as a function of the number of codewords was derived. There is a possibility that a linearly independent Plücker coordinates can be selected using specialized methods without necessary creating such columns using Gaussian elimination method. These selection methods would be studied further. This is because eliminating the Gaussian elimination process leads to reduction in complexity. The criteria for enumerating the basis would be expanded from success probability distribution to the probability of convergence or ergodicity.

In the creation of isomorphic map from the New Mersenne number transform, it was observed with the increase in the field, the degree of the leading monomials increased. This degraded the performance of the isomorphism to withstand key recovery attacks. In the future, the New Mersenne number isomorphism would be strengthened using a reduction process in order to reduce the degree of the monomials due to evaluation from functional composition. In addition, the effect of the butterfly operation on the key size will be a focus of further experimentation. This thesis, spent effort on the sparse solution of multivariate polynomials when the polynomials have been reduced to its ideal. In the future, improved variant of Gröbner basis computation would be studied extensively in order to create ideals that are solvable using linearization methods.

Bibliography

[1] Kelechi Chuwkunonyerem Emerole and Said Boussakta. "Isomorphism in Multivariate Cryptography using the New Mersenne Number Transform". In: *ICC 2022-IEEE International Conference on Communications*. IEEE. 2022, pp. 547–552.

[2] 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.

[3] Gorjan Alagic et al. "Status report on the second round of the NIST post-quantum cryptography standardization process". In: *US Department of Commerce, NIST* (2020).

[4] Christopher M Bishop, Markus Svensén, and Christopher KI Williams. "GTM: The generative topographic mapping". In: *Neural computation* 10.1 (1998), pp. 215–234.

[5] Divesh Aggarwal, Daniel Dadush, and Noah Stephens-Davidowitz. "Solving the Closest Vector Problem in 2ˆ n Time–The Discrete Gaussian Strikes Again!" In: *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*. IEEE. 2015, pp. 563–582.

[6] Miklós Ajtai, Ravi Kumar, and Dandapani Sivakumar. "A sieve algorithm for the shortest lattice vector problem". In: *Proceedings of the thirty-third annual ACM symposium on Theory of computing*. ACM. 2001, pp. 601–610.

[7] Arjen Klaas Lenstra, Hendrik Willem Lenstra, and László Lovász. "Factoring polynomials with rational coefficients". In: *Mathematische Annalen* 261.4 (1982), pp. 515–534.

[8] Florent Chabaud and Jacques Stern. "The cryptographic security of the syndrome decoding problem for rank distance codes". In: *International Conference on the Theory and Application of Cryptology and Information Security*. Springer. 1996, pp. 368–381.

[9] Nicolas Aragon et al. “A new algorithm for solving the rank syndrome decoding problem”. In: *2018 IEEE International Symposium on Information Theory (ISIT)*. IEEE. 2018, pp. 2421–2425.

[10] Philippe Gaborit, Olivier Ruatta, and Julien Schrek. “On the complexity of the rank syndrome decoding problem”. In: *IEEE Transactions on Information Theory* 62.2 (2015), pp. 1006–1019.

[11] Sebastian Franco, Daniele Galloni, and Alberto Mariotti. “Bipartite field theories, cluster algebras and the Grassmannian”. In: *Journal of Physics A: Mathematical and Theoretical* 47.47 (2014), p. 474004.

[12] Alexander Postnikov, David Speyer, and Lauren Williams. “Matching polytopes, toric geometry, and the totally non-negative Grassmannian”. In: *Journal of Algebraic Combinatorics* 30.2 (2009), pp. 173–191.

[13] Ramy H Gohary and Timothy N Davidson. “Noncoherent MIMO communication: Grassmannian constellations and efficient detection”. In: *IEEE Transactions on Information Theory* 55.3 (2009), pp. 1176–1205.

[14] Klaus Metsch. “A characterization of Grassmann graphs”. In: *European Journal of Combinatorics* 16.6 (1995), pp. 639–644.

[15] Eugene Prange. “The use of information sets in decoding cyclic codes”. In: *IRE Transactions on Information Theory* 8.5 (1962), pp. 5–9.

[16] Ghazal Kachigar and Jean-Pierre Tillich. “Quantum information set decoding algorithms”. In: *International Workshop on Post-Quantum Cryptography*. Springer. 2017, pp. 69–89.

[17] Nour-Eddine Lasmar and Yannick Berthoumieu. “Gaussian copula multivariate modeling for texture image retrieval using wavelet transforms”. In: *IEEE Transactions on Image Processing* 23.5 (2014), pp. 2246–2261.

[18] X Liu. “Copula of trivariate Rayleigh distribution with exponential correlation”. In: *Electronics Letters* 47.10 (2011), pp. 624–626.

[19] Hang Yu, Wayne Isaac T Uy, and Justin Dauwels. “Modeling spatial extremes via ensemble-of-trees of pairwise copulas”. In: *IEEE Transactions on Signal Processing* 65.3 (2016), pp. 571–586.

[20] Said Boussakta, Monir T Hamood, and Nick Rutter. “Generalized new mersenne number transforms”. In: *IEEE transactions on signal processing* 60.5 (2012), pp. 2640–2647.

[21] S Boussakta et al. “Number theoretic transforms of periodic structures and their applications”. In: *IEE Proceedings G (Electronic Circuits and Systems)*. Vol. 135. 2. IET. 1988, pp. 83–96.

[22] Jacques Patarin. “Asymmetric cryptography with a hidden monomial”. In: *Annual International Cryptology Conference*. Springer. 1996, pp. 45–60.

[23] Jacques Patarin. “Hidden fields equations (HFE) and isomorphisms of polynomials (IP): Two new families of asymmetric algorithms”. In: *International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 1996, pp. 33–48.

[24] Luk Bettale, Jean-Charles Faugere, and Ludovic Perret. “Hybrid approach for solving multivariate systems over finite fields”. In: *Journal of Mathematical Cryptology* 3.3 (2009), pp. 177–197.

[25] Yanbin Pan and Feng Zhang. “Solving low-density multiple subset sum problems with SVP oracle”. In: *Journal of Systems Science and Complexity* 29.1 (2016), pp. 228–242.

[26] Chris Peikert. “An efficient and parallel Gaussian sampler for lattices”. In: *Annual Cryptology Conference*. Springer. 2010, pp. 80–97.

[27] Zheng Wang and Cong Ling. “On the geometric ergodicity of Metropolis-Hastings algorithms for lattice Gaussian sampling”. In: *IEEE Transactions on Information Theory* 64.2 (2018), pp. 738–751.

[28] Carlos Aguilar-Melchor and Thomas Ricosset. “CDT-Based Gaussian Sampling: From Multi to Double Precision”. In: *IEEE Transactions on Computers* 67.11 (2018), pp. 1610–1621.

[29] Haijun Wang et al. “Single Image Super-Resolution Using Gaussian Process Regression With Dictionary-Based Sampling and Student Likelihood”. In: *IEEE Transactions on Image Processing* 26.7 (2017), pp. 3556–3568.

[30] Philip Klein. “Finding the closest lattice vector when it’s unusually close”. In: *Proceedings of the eleventh annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics. 2000, pp. 937–941.

[31] László Babai. “On Lovász’lattice reduction and the nearest lattice point problem”. In: *Combinatorica* 6.1 (1986), pp. 1–13.

[32] János Folláth. “Gaussian sampling in lattice based cryptography”. In: *Tatra Mountains Mathematical Publications* 60.1 (2014), pp. 1–23.

[33] Rafael Misoczki and Paulo SLM Barreto. "Compact McEliece keys from Goppa codes". In: *International Workshop on Selected Areas in Cryptography*. Springer. 2009, pp. 376–392.

[34] Chris Monico, Joachim Rosenthal, and Amin Shokrollahi. "Using low density parity check codes in the McEliece cryptosystem". In: *2000 IEEE International Symposium on Information Theory (Cat. No. 00CH37060)*. IEEE. 2000, p. 215.

[35] Marco Baldi et al. "Quasi-cyclic low-density parity-check codes in the mceliece cryptosystem". In: *2007 IEEE International Conference on Communications*. IEEE. 2007, pp. 951–956.

[36] Ernst M Gabidulin et al. "Reducible rank codes and their applications to cryptography". In: *IEEE Transactions on Information Theory* 49.12 (2003), pp. 3289–3293.

[37] Alexei V Ourivski and Thomas Johansson. "New technique for decoding codes in the rank metric and its cryptography applications". In: *Problems of Information Transmission* 38.3 (2002), pp. 237–246.

[38] Jean-Charles Faugere et al. "Fast quantum algorithm for solving multivariate quadratic equations". In: *arXiv preprint arXiv:1712.07211* (2017).

[39] Charles Bouillaguet, Pierre-Alain Fouque, and Amandine Véber. "Graph-theoretic algorithms for the "Isomorphism of Polynomials" problem". In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2013, pp. 211–227.

[40] Chengdong Tao et al. "Simple matrix scheme for encryption". In: *International Workshop on Post-Quantum Cryptography*. Springer. 2013, pp. 231–242.

[41] Tsutomu Matsumoto and Hideki Imai. "Public quadratic polynomial-tuples for efficient signature-verification and message-encryption". In: *Workshop on the Theory and Application of of Cryptographic Techniques*. Springer. 1988, pp. 419–453.

[42] Jacques Patarin, Nicolas Courtois, and Louis Goubin. "Quartz, 128-bit long digital signatures". In: *Cryptographers' Track at the RSA Conference*. Springer. 2001, pp. 282–297.

[43] Jintai Ding. "Gui-Algorithm Specification and Documentation". In: (2017), pp. 1–32.

[44] Netanel Raviv, Ben Langton, and Itzhak Tamo. "Multivariate Public Key Cryptosystem from Sidon Spaces." In: *Public Key Cryptography (1)*. 2021, pp. 242–265.

[45] Karina Mochetti and Ricardo Dahab. *Ideal lattice-based (H) IBE scheme*. Tech. rep. Technical Report IC-14-18, Institute of Computing, University of Campinas, 2014.

[46] James Howe et al. "Lattice-based encryption over standard lattices in hardware". In: *Proceedings of the 53rd Annual Design Automation Conference*. ACM. 2016, p. 162.

[47] Zecheng Wang, Xuemin Chen, and Pingshui Wang. "Adaptive-ID secure identity-based signature scheme from lattices in the standard model". In: *IEEE Access* 5 (2017), pp. 20791–20799.

[48] Zheng-She Liu. "QR methods of O (N) complexity in adaptive parameter estimation". In: *IEEE transactions on signal processing* 43.3 (1995), pp. 720–729.

[49] Xue Jiang et al. "Multipath channel estimation using fast least-squares algorithm". In: *2011 Third International Conference on Communications and Mobile Computing*. IEEE. 2011, pp. 433–436.

[50] Atsushi Takayasu and Yohei Watanabe. "Lattice-based revocable identity-based encryption with bounded decryption key exposure resistance". In: *Australasian Conference on Information Security and Privacy*. Springer. 2017, pp. 184–204.

[51] Daniele Micciancio and Chris Peikert. "Trapdoors for lattices: Simpler, tighter, faster, smaller". In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2012, pp. 700–718.

[52] AG D'yachkov et al. "Cover-free families and superimposed codes: constructions, bounds and applications to cryptography and group testing". In: *Proceedings. 2001 IEEE International Symposium on Information Theory (IEEE Cat. No. 01CH37252)*. IEEE. 2001, p. 117.

[53] Shantian Cheng and Juanyang Zhang. "Adaptive-ID secure revocable identity-based encryption from lattices via subset difference method". In: *International Conference on Information Security Practice and Experience*. Springer. 2015, pp. 283–297.

[54] Chunli Yang et al. "Hierarchical identity-based broadcast encryption scheme from LWE". In: *Journal of Communications and Networks* 16.3 (2014), pp. 258–263.

[55] Xiufeng Zhao and Xiang Wang. "An efficient identity-based signcryption from lattice". In: *International Journal of Security and Its Applications* 8.2 (2014), pp. 363–369.

[56] Fenghe Wang, ZhenHua Liu, and Chunxiao Wang. "Full secure identity-based encryption scheme with short public key size over lattices in the standard model". In: *International Journal of Computer Mathematics* 93.6 (2016), pp. 854–863.

[57] Vadim Lyubashevsky, Adriana Palacio, and Gil Segev. "Public-key cryptographic primitives provably as secure as subset sum". In: *Theory of Cryptography Conference*. Springer. 2010, pp. 382–400.

[58] Oded Regev. "On lattices, learning with errors, random linear codes, and cryptography". In: *Journal of the ACM (JACM)* 56.6 (2009), p. 34.

[59] Miklós Ajtai. "Generating hard instances of lattice problems". In: *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*. ACM. 1996, pp. 99–108.

[60] David Cash et al. "Bonsai trees, or how to delegate a lattice basis". In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2010, pp. 523–552.

[61] H Krishna and S Morgera. "Fast O (n) complexity algorithms for diagonal innovation matrices". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 32.6 (1984), pp. 1189–1194.

[62] Dan Boneh and Matt Franklin. "Identity-based encryption from the Weil pairing". In: *Annual international cryptology conference*. Springer. 2001, pp. 213–229.

[63] Mingwu Zhang et al. "Attribute-based hash proof system under learning-with-errors assumption in obfuscator-free and leakage-resilient environments". In: *IEEE Systems Journal* 11.2 (2015), pp. 1018–1026.

[64] Nabil Alkeilani Alkadri et al. "A Framework to Select Parameters for Lattice-Based Cryptography." In: *IACR Cryptology ePrint Archive* 2017 (2017), p. 615.

[65] Limin Zhou, Zhengming Hu, and Fengju Lv. "A simple lattice-based PKE scheme". In: *SpringerPlus* 5.1 (2016), p. 1627.

[66] Zoe L Jiang et al. "Lattice-based proxy signature scheme with reject sampling method". In: *2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*. IEEE. 2017, pp. 558–563.

[67] Philippe Gaborit et al. "RankSign: an efficient signature algorithm based on the rank metric". In: *International Workshop on Post-Quantum Cryptography*. Springer. 2014, pp. 88–107.

[68] Philippe Gaborit et al. “Low rank parity check codes and their application to cryptography”. In: *Proceedings of the Workshop on Coding and Cryptography WCC*. Vol. 2013. 2013.

[69] Donghoon Chang et al. “Revocable identity-based encryption from codes with rank metric”. In: *Cryptographers’ Track at the RSA Conference*. Springer. 2018, pp. 435–451.

[70] Adrien Hauteville and Jean-Pierre Tillich. “New algorithms for decoding in the rank metric and an attack on the LRPC cryptosystem”. In: *2015 IEEE International Symposium on Information Theory (ISIT)*. IEEE. 2015, pp. 2747–2751.

[71] Haruhiko Kaneko and Eiji Fujiwara. “A class of m-ary asymmetric symbol error correcting codes constructed by graph coloring”. In: *Proceedings. 2001 IEEE International Symposium on Information Theory (IEEE Cat. No. 01CH37252)*. IEEE. 2001, p. 225.

[72] Frank R Kschischang, Brendan J Frey, and H-A Loeliger. “Factor graphs and the sum-product algorithm”. In: *IEEE Transactions on information theory* 47.2 (2001), pp. 498–519.

[73] Azad Ravanshid, Lutz Lampe, and Johannes B Huber. “Dynamic decode-and-forward relaying using raptor codes”. In: *IEEE Transactions on Wireless Communications* 10.5 (2011), pp. 1569–1581.

[74] Zhong Cheng, Jeff Castura, and Yongyi Mao. “On the design of raptor codes for binary-input Gaussian channels”. In: *IEEE Transactions on Communications* 57.11 (2009), pp. 3269–3277.

[75] Amrit Kharel and Lei Cao. “Analysis and design of physical layer raptor codes”. In: *IEEE Communications Letters* 22.3 (2017), pp. 450–453.

[76] Thomas P Minka. “Expectation propagation for approximate Bayesian inference”. In: *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc. 2001, pp. 362–369.

[77] Francisco Lázaro, Gianluigi Liva, and Gerhard Bauch. “Inactivation decoding of LT and Raptor codes: Analysis and code design”. In: *IEEE Transactions on Communications* 65.10 (2017), pp. 4114–4127.

[78] Sachini Jayasooriya et al. “Analysis and design of Raptor codes using a multi-edge framework”. In: *IEEE Transactions on Communications* 65.12 (2017), pp. 5123–5136.

[79] Kaveh Mahdaviani, Masoud Ardakani, and Chintha Tellambura. “On Raptor code design for inactivation decoding”. In: *IEEE Transactions on Communications* 60.9 (2012), pp. 2377–2381.

[80] Mahyar Shirvanimoghaddam and Sarah Johnson. “Raptor codes in the low SNR regime”. In: *IEEE Transactions on Communications* 64.11 (2016), pp. 4449–4460.

[81] Guy Even and Nissim Halabi. “On decoding irregular Tanner codes with local-optimality guarantees”. In: *CoRR* (2011).

[82] Noga Alon et al. “Construction of asymptotically good low-rate error-correcting codes through pseudo-random graphs”. In: *IEEE Transactions on information theory* 38.2 (1992), pp. 509–516.

[83] Amir H Banihashemi, FR Kschischang, and P Glenn Gulak. “On Tanner graphs of lattices and codes”. In: *Proceedings. 1998 IEEE International Symposium on Information Theory (Cat. No. 98CH36252)*. IEEE. 1998, p. 115.

[84] Veria Havary-Nassab, Shahram Shahbazpanahi, and Ali Grami. “Optimal distributed beamforming for two-way relay networks”. In: *IEEE Transactions on Signal Processing* 58.3 (2009), pp. 1238–1250.

[85] Mehdi Hassani. “Enumeration by e”. In: *Modern Discrete Mathematics and Analysis*. Springer, 2018, pp. 227–233.

[86] Sven Müelich et al. “An alternative decoding method for Gabidulin codes in characteristic zero”. In: *2016 IEEE International Symposium on Information Theory (ISIT)*. IEEE. 2016, pp. 2549–2553.

[87] Joachim von zur Gathen. “Irreducibility of multivariate polynomials”. In: *Journal of Computer and System Sciences* 31.2 (1985), pp. 225–264.

[88] Aviad Kipnis, Jacques Patarin, and Louis Goubin. “Unbalanced oil and vinegar signature schemes”. In: *International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 1999, pp. 206–222.

[89] Jean-Charles Faugere. “A new efficient algorithm for computing Gröbner bases (F4)”. In: *Journal of pure and applied algebra* 139.1-3 (1999), pp. 61–88.

[90] Nicolas T Courtois. “Higher order correlation attacks, XL algorithm and cryptanalysis of Toyocrypt”. In: *International Conference on Information Security and Cryptology*. Springer. 2002, pp. 182–199.

[91] Antoine Joux and Vanessa Vitse. “A crossbred algorithm for solving Boolean polynomial systems”. In: *International Conference on Number-Theoretic Methods in Cryptology*. Springer. 2017, pp. 3–21.

[92] Olaf Manz and Thomas R Wolf. *Representations of solvable groups*. Vol. 185. Cambridge University Press, 1993.

[93] Lih-Chung Wang et al. "A "medium-field" multivariate public-key encryption scheme". In: *Cryptographers' Track at the RSA Conference*. Springer. 2006, pp. 132–149.

[94] Jaiberth Porras, John Baena, and Jintai Ding. "ZHFE, a new multivariate public key encryption scheme". In: *International workshop on post-quantum cryptography*. Springer. 2014, pp. 229–245.

[95] Bo-Yin Yang and Jiun-Ming Chen. "Building secure tame-like multivariate public-key cryptosystems: The new TTS". In: *Australasian Conference on Information Security and Privacy*. Springer. 2005, pp. 518–531.

[96] Daniel J Bernstein and Bo-Yin Yang. "Asymptotically faster quantum algorithms to solve multivariate quadratic equations". In: *International Conference on Post-Quantum Cryptography*. Springer. 2018, pp. 487–506.

[97] Vasco Manquinho, Ruben Martins, and Inês Lynce. "Improving unsatisfiability-based algorithms for boolean optimization". In: *International conference on theory and applications of satisfiability testing*. Springer. 2010, pp. 181–193.

[98] Benjamin Pring. "Exploiting Preprocessing for Quantum Search to Break Parameters for

\mathcal {MQ}

Cryptosystems

". In: *International Workshop on the Arithmetic of Finite Fields*. Springer. 2018, pp. 291–307.

[99] Nicolas Courtois et al. "Efficient algorithms for solving overdefined systems of multivariate polynomial equations". In: *International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2000, pp. 392–407.

[100] Aarti Dadheech. "Preventing Information Leakage from Encoded Data in Lattice Based Cryptography". In: *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE. 2018, pp. 1952–1955.

[101] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. "On ideal lattices and learning with errors over rings". In: *Journal of the ACM (JACM)* 60.6 (2013), p. 43.

[102] Vadim Lyubashevsky and Daniel Wichs. “Simple lattice trapdoor sampling from a broad class of distributions”. In: *IACR International Workshop on Public Key Cryptography*. Springer. 2015, pp. 716–730.

[103] Carlos Aguilar-Melchor, Martin R Albrecht, and Thomas Ricosset. “Sampling from arbitrary centered discrete gaussians for lattice-based cryptography”. In: *International Conference on Applied Cryptography and Network Security*. Springer. 2017, pp. 3–19.

[104] Léo Ducas and Thomas Prest. “A hybrid Gaussian sampler for lattices over rings”. In: *IACR Cryptology ePrint Archive* (2015), p. 660.

[105] Zhenguo Gao et al. “Probability density function of the Euclidean distance between node pairs in rectangular random graphs and its applications in MANETs”. In: *CHINESE JOURNAL OF ELECTRONICS* 15.3 (2006), p. 521.

[106] Markku-Juhani O Saarinen. “Gaussian Sampling Precision and Information Leakage in Lattice Cryptography.” In: *IACR Cryptology ePrint Archive* 2015 (2015), p. 953.

[107] Claus Peter Schnorr. “Progress on LLL and lattice reduction”. In: *The LLL Algorithm*. Springer, 2009, pp. 145–178.

[108] Henning Vetter et al. “Fixed complexity LLL algorithm”. In: *IEEE Transactions on signal processing* 57.4 (2009), pp. 1634–1637.

[109] Tadashi Fujino. “Gram-Schmidt combined LLL lattice-reduction aided detection in MIMO systems”. In: *REV Journal on Electronics and Communications* 1.2 (2011).

[110] Hongfei Zhu et al. “An identity-based proxy signature on NTRU lattice”. In: *Chinese Journal of Electronics* 27.2 (2018), pp. 297–303.

[111] Léo Ducas and Phong Q Nguyen. “Faster Gaussian lattice sampling using lazy floating-point arithmetic”. In: *International Conference on the Theory and Application of Cryptology and Information Security*. Springer. 2012, pp. 415–432.

[112] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. “Trapdoors for hard lattices and new cryptographic constructions”. In: *Proceedings of the fortieth annual ACM symposium on Theory of computing*. 2008, pp. 197–206.

[113] Vadim Lyubashevsky and Thomas Prest. “Quadratic time, linear space algorithms for Gram-Schmidt orthogonalization and Gaussian sampling in structured lattices”. In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2015, pp. 789–815.

[114] Angshuman Karmakar et al. “Constant-time discrete gaussian sampling”. In: *IEEE Transactions on Computers* 67.11 (2018), pp. 1561–1571.

[115] Chaohui Du and Guoqiang Bai. “Towards efficient discrete Gaussian sampling for lattice-based cryptography”. In: *2015 25th International Conference on Field Programmable Logic and Applications (FPL)*. IEEE. 2015, pp. 1–6.

[116] Nicholas Genise and Daniele Micciancio. “Faster gaussian sampling for trapdoor lattices with arbitrary modulus”. In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2018, pp. 174–203.

[117] Daniele Micciancio and Michael Walter. “Gaussian sampling over the integers: Efficient, generic, constant-time”. In: *Annual International Cryptology Conference*. Springer. 2017, pp. 455–485.

[118] Nicolas Gama and Phong Q Nguyen. “Predicting lattice reduction”. In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2008, pp. 31–51.

[119] Luc Giraud, Julien Langou, and Miroslav Rozloznik. “The loss of orthogonality in the Gram-Schmidt orthogonalization process”. In: *Computers & Mathematics with Applications* 50.7 (2005), pp. 1069–1075.

[120] Charles M Werneth et al. “Numerical gram–schmidt orthonormalization”. In: *European Journal of Physics* 31.3 (2010), p. 693.

[121] Jian Zhao, Haiying Gao, and Junqi Zhang. “Attribute-based encryption for circuits on lattices”. In: *Tsinghua Science and Technology* 19.5 (2014), pp. 463–469.

[122] Daniele Micciancio and Oded Regev. “Worst-case to average-case reductions based on Gaussian measures”. In: *SIAM Journal on Computing* 37.1 (2007), pp. 267–302.

[123] Thomas Pöppelmann and Tim Güneysu. “Area optimization of lightweight lattice-based encryption on reconfigurable hardware”. In: *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE. 2014, pp. 2796–2799.

[124] Xiaolin Wu, Philip A Chou, and Xiaohui Xue. “Minimum conditional entropy context quantization”. In: *IEEE International Symposium on Information Theory*. Citeseer. 2000, pp. 43–43.

[125] David W Scott. *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons, 2015.

[126] Mingsheng Shang et al. “Randomized latent factor model for high-dimensional and sparse matrices from industrial applications”. In: *IEEE/CAA Journal of Automatica Sinica* 6.1 (2019), pp. 131–141.

[127] Cem Örnek and Elif Vural. "Nonlinear supervised dimensionality reduction via smooth regular embeddings". In: *Pattern Recognition* 87 (2019), pp. 55–66.

[128] Jun Guo et al. "Discriminative unsupervised 2D dimensionality reduction with graph embedding". In: *Multimedia Tools and Applications* 77.3 (2018), pp. 3189–3207.

[129] Mehrtash Harandi, Mathieu Salzmann, and Richard Hartley. "Dimensionality reduction on SPD manifolds: The emergence of geometry-aware methods". In: *IEEE transactions on pattern analysis and machine intelligence* 40.1 (2018), pp. 48–62.

[130] Qi Mao et al. "Principal graph and structure learning based on reversed graph embedding". In: *IEEE transactions on pattern analysis and machine intelligence* 39.11 (2017), pp. 2227–2241.

[131] Theodoros Iliou and Christos-Nikolaos Anagnostopoulos. "SVM-MLP-PNN classifiers on speech emotion recognition field-A comparative study". In: *2010 Fifth International Conference on Digital Telecommunications*. IEEE. 2010, pp. 1–6.

[132] Johan AK Suykens, Lukas Lukas, and Joos Vandewalle. "Sparse approximation using least squares support vector machines". In: *2000 IEEE International Symposium on Circuits and Systems. Emerging Technologies for the 21st Century. Proceedings (IEEE Cat No. 00CH36353)*. Vol. 2. IEEE. 2000, pp. 757–760.

[133] Nicolas Gillis and Stephen A Vavasis. "On the complexity of robust pca and 1-norm low-rank matrix approximation". In: *Mathematics of Operations Research* 43.4 (2018), pp. 1072–1084.

[134] Cho-Jui Hsieh et al. "QUIC: quadratic approximation for sparse inverse covariance estimation". In: *The Journal of Machine Learning Research* 15.1 (2014), pp. 2911–2947.

[135] Symeon Chouvardas, Konstantinos Slavakis, and Sergios Theodoridis. "Trading off complexity with communication costs in distributed adaptive learning via Krylov subspaces for dimensionality reduction". In: *IEEE Journal of Selected Topics in Signal Processing* 7.2 (2013), pp. 257–273.

[136] Arun Kumar Singh, Petros Elia, and Joakim Jaldén. "Achieving a vanishing SNR gap to exact lattice decoding at a subexponential complexity". In: *IEEE Transactions on Information Theory* 58.6 (2012), pp. 3692–3707.

[137] Shahriar Shahabuddin et al. "A customized lattice reduction multiprocessor for MIMO detection". In: *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE. 2015, pp. 2976–2979.

[138] José Carlos Marinello and Taufik Abrão. "Lattice reduction aided detector for dense MIMO via ant colony optimization". In: *2013 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE. 2013, pp. 2839–2844.

[139] Svante Bergman. "Bit loading and precoding for MIMO communication systems". PhD thesis. KTH, 2009.

[140] YuPu Hu et al. "Gaussian sampling of lattices for cryptographic applications". In: *Science China Information Sciences* 57.7 (2014), pp. 1–8.

[141] Shi Bai et al. "Improved security proofs in lattice-based cryptography: using the Rényi divergence rather than the statistical distance". In: *Journal of Cryptology* 31.2 (2018), pp. 610–640.

[142] Léo Ducas et al. "Lattice signatures and bimodal Gaussians". In: *Annual Cryptology Conference*. Springer. 2013, pp. 40–56.

[143] Daniel J Bernstein, Tanja Lange, and Christiane Peters. "Smaller decoding exponents: ball-collision decoding". In: *Annual Cryptology Conference*. Springer. 2011, pp. 743–760.

[144] Elwyn Berlekamp, Robert McEliece, and Henk Van Tilborg. "On the inherent intractability of certain coding problems (corresp.)" In: *IEEE Transactions on Information Theory* 24.3 (1978), pp. 384–386.

[145] Robert J McEliece. "A public-key cryptosystem based on algebraic". In: *Coding Thv* 4244 (1978), pp. 114–116.

[146] Jacques Stern. "A method for finding codewords of small weight". In: *International Colloquium on Coding Theory and Applications*. Springer. 1988, pp. 106–113.

[147] Tuvi Etzion and Hui Zhang. "Grassmannian codes with new distance measures for network coding". In: *IEEE Transactions on Information Theory* 65.7 (2019), pp. 4131–4142.

[148] Charles T Ryan and Kevin M Ryan. "The minimum weight of the Grassmann codes C (k, n)". In: *Discrete applied mathematics* 28.2 (1990), pp. 149–156.

[149] Tuvi Etzion and Natalia Silberstein. "Codes and designs related to lifted MRD codes". In: *IEEE Transactions on Information Theory* 59.2 (2012), pp. 1004–1017.

[150] R Tanner. "A recursive approach to low complexity codes". In: *IEEE Transactions on information theory* 27.5 (1981), pp. 533–547.

[151] Ralf Koetter and Frank R Kschischang. "Coding for errors and erasures in random network coding". In: *IEEE Transactions on Information theory* 54.8 (2008), pp. 3579–3591.

[152] Rom Rubenovich Varshamov. "The evaluation of signals in codes with correction of errors". In: *Doklady Akademii Nauk*. Vol. 117. 5. Russian Academy of Sciences. 1957, pp. 739–741.

[153] George E Andrews. *q-Series: Their Development and Application in Analysis, Number Theory, Combinatorics, Physics and Computer Algebra: Their Development and Application in Analysis, Number Theory, Combinatorics, Physics, and Computer Algebra*. 66. American Mathematical Soc., 1986.

[154] EM Gabidulin and NI Pilipchuk. "Subspace Network Codes with Large Cardinality". In: *2015 International Conference on Engineering and Telecommunication (EnT)*. IEEE. 2015, pp. 10–13.

[155] Qian Guo et al. "Some cryptanalytic and coding-theoretic applications of a soft stern algorithm". In: *Advances in Mathematics of Communications* 13.4 (2019), p. 559.

[156] Ward Beullens. *Not enough LESS: An improved algorithm for solving Code Equivalence Problems over Fq*. Tech. rep.

[157] Daniel J Bernstein. "Introduction to post-quantum cryptography". In: *Post-quantum cryptography*. Springer, 2009, pp. 1–14.

[158] Philippe Gaborit et al. "Identity-based encryption from codes with rank metric". In: *Annual International Cryptology Conference*. Springer. 2017, pp. 194–224.

[159] Carlos Aguilar Melchor et al. "A new efficient threshold ring signature scheme based on coding theory". In: *IEEE Transactions on Information Theory* 57.7 (2011), pp. 4833–4842.

[160] Terry Lau and Chik Tan. "A new technique in rank metric code-based encryption". In: *Cryptography* 2.4 (2018), p. 32.

[161] Pongpol Ruankong and Songkiat Sumetkijakan. "Essential closures and supports of multivariate copulas". In: *International journal of approximate reasoning* 54.6 (2013), pp. 762–768.

[162] Shanxiang Lyu and Cong Ling. "Hybrid vector perturbation precoding: The blessing of approximate message passing". In: *IEEE Transactions on Signal Processing* 67.1 (2018), pp. 178–193.

[163] Charles Bouillaguet et al. "A family of weak keys in HFE and the corresponding practical key-recovery". In: *Journal of Mathematical Cryptology* 5.3-4 (2012), pp. 247–275.

[164] Dung Hoang Duong, Albrecht Petzoldt, and Tsuyoshi Takagi. "Reducing the key size of the SRP encryption scheme". In: *Australasian Conference on Information Security and Privacy*. Springer. 2016, pp. 427–434.

[165] Christopher Wolf and Bart Preneel. "Taxonomy of Public Key Schemes based on the problem of Multivariate Quadratic equations." In: *IACR Cryptology ePrint Archive* 2005 (2005), p. 77.

[166] Ling Yang and Xianhui Lu. "An efficient dispersal storage scheme based on Ring-LWE and NTT". In: *2017 12th Asia Joint Conference on Information Security (AsiaJCIS)*. IEEE. 2017, pp. 23–30.

[167] Soo-Chang Pei, Chia-Chang Wen, and Jian-Jiun Ding. "Closed-form orthogonal number theoretic transform eigenvectors and the fast fractional NTT". In: *IEEE transactions on signal processing* 59.5 (2011), pp. 2124–2135.

[168] Gregor Seiler. "Faster AVX2 optimized NTT multiplication for Ring-LWE lattice cryptography." In: *IACR Cryptology ePrint Archive* 2018 (2018), p. 39.

[169] M Perz. "Thomas koshy," elementary number theory with applications", 2002". In: *Smarandache Notions Journal* 13 (2002), pp. 284–285.

[170] Jintai Ding et al. "High order linearization equation (hole) attack on multivariate public key cryptosystems". In: *International Workshop on Public Key Cryptography*. Springer. 2007, pp. 233–248.

[171] Aviad Kipnis and Adi Shamir. "Cryptanalysis of the HFE public key cryptosystem by relinearization". In: *Annual International Cryptology Conference*. Springer. 1999, pp. 19–30.

[172] Rachid El Bansarkhani, Mohamed Saied Emam Mohamed, and Albrecht Petzoldt. "MQSAS-a multivariate sequential aggregate signature scheme". In: *International Conference on Information Security*. Springer. 2016, pp. 426–439.

[173] Willi Geiselmann, Willi Meier, and Rainer Steinwandt. "An attack on the isomorphisms of polynomials problem with one secret". In: *International Journal of Information Security* 2.1 (2003), pp. 59–64.

[174] Carl G Ponder. "Parallel multiplication and powering of polynomials". In: *Journal of symbolic computation* 11.4 (1991), pp. 307–320.

[175] Jintai Ding et al. "Improved cryptanalysis of hfev-via projection". In: *International Conference on Post-Quantum Cryptography*. Springer. 2018, pp. 375–395.

[176] M Vanwormhoudt. “On number theoretic Fourier transforms in residue class rings”. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 25.6 (1977), pp. 585–586.

[177] Yasuhiko Ikematsu et al. “HFERP-a new multivariate encryption scheme”. In: *International Conference on Post-Quantum Cryptography*. Springer. 2018, pp. 396–416.

[178] Patrick Longa and Michael Naehrig. “Speeding up the number theoretic transform for faster ideal lattice-based cryptography”. In: *International Conference on Cryptology and Network Security*. Springer. 2016, pp. 124–139.

[179] Albrecht Petzoldt et al. “Design principles for HFEv-based multivariate signature schemes”. In: *International Conference on the Theory and Application of Cryptology and Information Security*. Springer. 2015, pp. 311–334.

[180] The OpenSSL Project. “OpenSSL: The Open Source toolkit for SSL/TLS”. `www.openssl.org`. 2003.

[181] Y Chen and C Huang. “Efficient Operations In Large Finite Fields For Elliptic Curve Cryptographic”. In: *International Journal of Engineering Technologies and Management Research* 7.6 (2020), pp. 141–151.

[182] WA Stein et al. *SageMathCloud Software*. 2015.

[183] Jean-Charles Faugère. “FGb: a library for computing Gröbner bases”. In: *International Congress on Mathematical Software*. Springer. 2010, pp. 84–87.

[184] Jean-Charles Faugere and Ludovic Perret. “An efficient algorithm for decomposing multivariate polynomials and its applications to cryptography”. In: *Journal of Symbolic Computation* 44.12 (2009), pp. 1676–1689.

[185] Alessio Caminata and Elisa Gorla. “Solving multivariate polynomial systems and an invariant from commutative algebra”. In: *arXiv preprint arXiv:1706.06319* (2017).

[186] Christina Boura and Anne Canteaut. “On the Influence of the Algebraic Degree of F^{-1} on the Algebraic Degree of $G\backslash$circ F”. In: *IEEE Transactions on Information Theory* 59.1 (2012), pp. 691–702.

[187] Zhi-Quan Luo and Paul Tseng. “On the convergence of the coordinate descent method for convex differentiable minimization”. In: *Journal of Optimization Theory and Applications* 72.1 (1992), pp. 7–35.

[188] Philip Wolfe. “Convergence conditions for ascent methods”. In: *SIAM review* 11.2 (1969), pp. 226–235.

[189] J Mercer. *Functions of positive and negative type and their connection with the theory of integral equations, Philosophical Transsaction of the Royal Society of London, Ser.* 1909.

[190] Sanjeev Arora et al. "A convergence analysis of gradient descent for deep linear neural networks". In: *arXiv preprint arXiv:1810.02281* (2018).

[191] Nicolas Gama, Phong Q Nguyen, and Oded Regev. "Lattice enumeration using extreme pruning". In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2010, pp. 257–278.

[192] Claus-Peter Schnorr and Martin Euchner. "Lattice basis reduction: Improved practical algorithms and solving subset sum problems". In: *Mathematical programming* 66.1 (1994), pp. 181–199.

[193] Milton Abramowitz and Irene A Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Vol. 55. Courier Corporation, 1965.

Appendix

6.A Appendix A

Proof. The expectation of the BIAWGN channel that defines the distribution between the intermediate vertex and output vertex is dependent on the expansion of the mean of the distribution to its central of origin where the mean is given by $\mu_n = \frac{x-\mu}{4\mu}.e^{\frac{-(x-\mu)^2}{\sqrt{4\mu}}}dx$

$$E(BIAWGN)(\sigma)) = \frac{1}{2\sqrt{\pi\mu}}\int_{-\infty}^{\infty}\zeta(x^{(k)})e^{\frac{-(x-\mu)^2}{4\mu}}dx \tag{6.2}$$

For simplicity we take $P\zeta(x^{(k)}) = X$, then we have

$$\begin{aligned}
&= \frac{1}{2\sqrt{\pi\mu}}\left(\int_{-\infty}^{\infty}\frac{e^X - e^{-X}}{e^X + e^{-X}}e^{\frac{-(x-\mu)^2}{4\mu}}dx - \right. \\
&\qquad \left.\int_{-\infty}^{\infty}\frac{e^{-X}}{e^X + e^{-X}}.e^{\frac{-(x-\mu)^2}{4\mu}}dx\right) \\
&= \frac{1}{2\sqrt{\pi\mu}}\left(\int_{-\infty}^{\infty}\frac{1}{1 + e^{-2X}}e^{\frac{-(x-\mu)^2}{4\mu}}dx - \right. \\
&\qquad \left.\int_{-\infty}^{\infty}\frac{1}{e^{2X} + 1}.e^{\frac{-(x-\mu)^2}{4\mu}}dx\right) \\
&= \sqrt{4\mu}\int_{-\infty}^{\infty}\frac{1}{1 + e^{-2(\sqrt{4\mu}(\frac{x-\mu}{\sqrt{4\mu}})+\mu}}e^{\frac{-(x-\mu)^2}{\sqrt{4\mu}}}dx + \\
&\quad \sqrt{4\mu}\int_{-\infty}^{\infty}\frac{1}{1 + e^{-2(\sqrt{4\mu(\frac{x-\mu}{\sqrt{4\mu}})}+\mu}}e^{\frac{-(x-\mu)^2}{\sqrt{4\mu}}}dx \\
&= \int_{-\infty}^{+\infty} ln(e^{2\sqrt{4\mu}(\frac{x-\mu}{\sqrt{4\mu}})+\mu} + \frac{x-\mu}{4\mu}.e^{\frac{-(x-\mu)^2}{\sqrt{4\mu}}}dx
\end{aligned} \tag{6.3}$$

taking the *nth* central moment about the mean of the distribution

$$\mu_n = \frac{x-\mu}{4\mu}.e^{\frac{-(x-\mu)^2}{\sqrt{4\mu}}}dx \tag{6.4}$$

then expanding the central moment to its centroid of origin[193], we have

$$= E_{in} + \sum_{j=0}^{n} \binom{n}{j} (-1)^{n-j} \mu_j^{'} \mu^{n-j} \tag{6.5}$$

where E_{in} is the decoding error probability and
$\mu_j^{'} = \left(1 - \frac{\alpha}{n}\right)^{n-l}$ □

6.B Appendix B

Proof. Given the probability of generating an output vertex as follows $\sum_{\alpha=1}^{d} \binom{n}{l} \binom{\alpha}{n}^{d} \left(1 - \frac{\alpha}{n}\right)^{n-l} x^d$ where $\binom{n}{l}$ is the parametrization of the subspace, in which n is the dimension of the vector space and l is the initial dimension of the output subspace and $\binom{\alpha}{n}$ is the parametrization of the received syndrome space . This follows that

$$\forall d \geq 2 \tag{6.6}$$

$$= \sum_{\alpha=1}^{d} (l(l-1)+l) \binom{n}{l} \binom{\alpha}{n}^{d} \left(1-\frac{\alpha}{n}\right)^{n-l}$$

$$= \sum_{\alpha=1}^{d} (l(l-1)) \binom{n}{l} \binom{\alpha}{n}^{d} \left(1-\frac{\alpha}{n}\right)^{n-l} + \sum_{\alpha=1}^{d} l \binom{n}{l} \binom{\alpha}{n}^{d} \left(1-\frac{\alpha}{n}\right)^{n-l}$$

$$= n(n-1) \binom{\alpha}{n}^{2} \sum_{\alpha=2}^{d} \binom{n-2}{l-2} \binom{\alpha}{n}^{l-2} \left(1-\frac{\alpha}{n}\right)^{n-2-(n-l)} + \sum_{\alpha=1}^{d} l \binom{n}{l} \binom{\alpha}{n}^{d} \left(1-\frac{\alpha}{n}\right)^{n-l}$$

$$\forall d \in \{1, \ldots, j\} \tag{6.7}$$

$$= n(n-1) \binom{\alpha}{n}^{2} \sum_{j=0}^{d-2} \binom{n-2}{j} \binom{\alpha}{n}^{j} \left(1-\frac{\alpha}{n}\right)^{n-2-j} + \sum_{\alpha=1}^{d} l \binom{n}{l} \binom{n}{l} \left(1-\frac{\alpha}{n}\right)^{n-l}$$

$$= n(n-1) \binom{\alpha}{n}^{2} + \sum_{\alpha=1}^{d} l \binom{n}{l} \binom{\alpha}{n}^{d} \left(1-\frac{\alpha}{n}\right)^{n-l} + \sum_{\alpha=1}^{d} l \binom{n}{l} \binom{\alpha}{n}^{d} \left(1-\frac{\alpha}{n}\right)^{n-l} \left(\left(n-1\right)\frac{\alpha}{n}+1\right)$$

$$= \sum_{\alpha=1}^{d} l \binom{n}{l} \binom{\alpha}{n}^{d} \left(1-\frac{\alpha}{n}\right)^{n-l} \left(\left(n-1\right)\frac{\alpha}{n}+1\right) \left(\sum_{\alpha=1}^{d} l \binom{n}{l} \binom{\alpha}{n}^{d} - \left(\frac{\alpha}{n}\right)+1\right)$$

$$= \Delta^{k} \alpha_{d}$$

where $\Delta^{k}\alpha_{d}$ is the discrete analog. □

6.C Appendix C

Proof. We start with a basis for U, $B_1 = (e_1, \ldots, e_m)$, picking randomly linearly independent vector $x_{U_i} \in U$. Then search for a coordinate of x_{U_i} and replace to pro-

duce a new basis for U after repeated procedures to give $B_1 = e_1^{'}, \dots e_m^{'}, x_{U_1}, \dots x_{U_k}$ and update count as

$$\texttt{Count}_U = \prod_{k=0}^{U_i - 1} q^k = \sum_{k=0}^{U_i} q^{\frac{k-1}{2}} \binom{n}{k}_q . \tag{6.8}$$

Then the same process follows for V as we start also start with a basis, $B_2 = (f_1, \dots, f_m)$. Then we pick random linearly independent vectors $y_{V_i} \in V$ and search for coordinate of y_{V_i} and replace to produce a new basis for V after repeated procedures to give $B_2 = f_1^{'}, \dots f_m^{'}, y_{V_1}, \dots y_{V_k}$ and update count as

$$\texttt{Count}_V = \prod_{k=0}^{V_i - 1} q^k - q^{k-r} = \sum_{k=0}^{V_i} q^{\frac{k(k-r)}{2}} \binom{k}{r}_q \tag{6.9}$$

Then, finally we start with a basis for $U \cap V$, $B_3 = (g_1, \dots, g_m)$, then we pick random linearly independent vector $z_i \in U \cap V$ to produce a new basis after repeated procedures $B_{3'} = (g_1^{'}, \dots g_m^{'}, x_{U_1}, \dots, x_{U_k}$ and $B_{3''} = (g_1^{'}, \dots g_m^{'}, y_{V_1}, \dots, y_{V_k})$. Sampling an integer $l_i \in L$ where $L = \texttt{Vect}(x_U)$ and $p_i \in P$ where $P = \texttt{Vect}(y_V)$ and updating the count as

$$\texttt{Count}_* = \prod_{k=0}^{U_i - V_i - 1} q^k - q^{k-r+t} - q^{k-r+p} = \sum_{k=0}^{U_i - V_i - 1} q^{\frac{k(k-r)}{2}} \begin{bmatrix} n-r \\ k-r \end{bmatrix}_q . \begin{bmatrix} r \\ k-t \end{bmatrix}_q . \begin{bmatrix} k \\ r \end{bmatrix}_q \tag{6.10}$$

From the total of the Counts, $\texttt{Count} = \texttt{Count}_U + \texttt{Count}_V + \texttt{Count}_*$, we can compute the bounds. It follows that $U = \texttt{span}\{g_i, x_{U_i}\}$, $V = span\{g_i, y_{V_i}\}$ and $U \cap V = \texttt{span}\{g_i\}$ □

6.D Appendix D

Proof.

$$\int_0^\infty C_1(1 - u_{1n})(1 - u_{2n})du_1 \int_0^\infty C_2(1 - u_{1n})u_{2n}du_2 + \tag{6.11}$$
$$\int_0^\infty C_3 u_{1n}(1 - u_{2n})du_1 \int_0^\infty C_4(u_{1n})(u_{2n})du_2$$

taking partial derivatives and sampling the copulas from a uniform distribution parameterized by the coordinates

$$u_{1n}[\ \frac{\partial C_3(u_{1n})}{\partial u} - \frac{\partial C_1(u_{2n})}{\partial u}]\ du_1 + \tag{6.12}$$
$$u_{2n}[\ \frac{\partial C_2(u_{1n})}{\partial u} - \frac{\partial C_1(u_{2n})}{\partial u}]\ du_2 +$$
$$u_{1n}u_{2n}[\ \frac{\partial C_1(u_{1n})}{\partial u} + \frac{\partial C_4(u_{2n})}{\partial u}]\ - \frac{\partial C_2(u_{1n})}{\partial u} -$$
$$\frac{\partial C_3(u_{2n})}{\partial u} du_1 du_2$$

taking Equation (24) and normalizing the other terms we have

$$\int \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{-\rho^2(x_1^2+x_2^2) - 2\rho x_1 x_2}{2(1-\rho^2)}\right) dx \tag{6.13}$$

integrating the exponential term

$$\int \exp\left(\frac{-\rho^2(x_1^2+x_2^2) - 2\rho x_1 x_2}{2(1-\rho^2)}\right) dx \tag{6.14}$$

$$\frac{\partial C(u,v)}{\partial u} = \frac{dx_2}{du}\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\rho x_2)^2}{2}\frac{1}{x_2}\frac{x_1-\rho x_2}{\sqrt{(1-\rho^2)}}\right) = \tag{6.15}$$

$$\sqrt{2\pi}\exp\left(\frac{2}{(\rho x_2)^2}\right)\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\rho x_2)^2}{2}\frac{1}{x_2}\frac{x_1-\rho x_2}{\sqrt{(1-\rho^2)}}\right) = \frac{1}{x_2}\frac{x_1-\rho x_2}{\sqrt{(1-\rho^2)}} \tag{6.16}$$

$$\frac{\partial^2}{\partial u \partial v} C(u,v) = \frac{1}{\sqrt{1-\rho^2}}\frac{dx_1}{dv}\frac{1}{\sqrt{2\pi}}\frac{x_1-\rho x_2}{\sqrt{1-\rho^2}} = \tag{6.17}$$

$$\frac{1}{\sqrt{1-\rho^2}}\left(\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\rho x_1)^2}{2}\frac{1}{\sqrt{2\pi}}\left(\frac{-x_2 + 2\rho x_2 x_1 - \rho^2 x_2^2}{2(1-\rho^2)}\right)\right) = \tag{6.18}$$

$$\frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{(\rho x_1)^2)}{2}\frac{-x_1^2 + 2\rho x_2 x_1 - \rho^2 x_2^2}{2}\right) = \tag{6.19}$$

$$\frac{1}{\sqrt{1-\rho^2}} \exp\left(\frac{-\rho^2 x_1^2 + 2\rho x_2 x}{2(1-\rho^2)}\right). \tag{6.20}$$

□