

The Hardness of LWE and Ring-LWE: A Survey

David Balbás¹

IMDEA Software Institute
Universidad Politécnica de Madrid

`david.balbas@imdea.org`

8th October 2021

¹Most of this work was carried out as a student at KTH Royal Institute of Technology, Stockholm.

Abstract

The Learning with Errors (LWE) problem consists of distinguishing linear equations with noise from uniformly sampled values. LWE enjoys a hardness reduction from worst-case lattice problems, which are believed to be hard for classical and quantum computers. Besides, LWE allows for the construction of a large variety of cryptographic schemes, including fully-homomorphic encryption and attribute-based cryptosystems. Unfortunately, LWE requires large key sizes and computation times. To improve efficiency, Ring-LWE replaces linear equations with noisy ring products. Nowadays, Ring-LWE and its variants are frequently used in the construction of post-quantum secure cryptosystems.

In this survey, we give an overview of the hardness results for LWE and Ring-LWE, aiming to connect both problems and to provide good intuition to the reader. We present a proof of the strongest hardness result for Ring-LWE available in the literature, which is a reduction from ideal lattice problems to its decision form. We start by introducing both Ring-LWE and LWE and their mathematical foundations, focusing on lattices and algebraic number theory. Then, we sketch the classical hardness proof for LWE and extend the proof techniques to the ring case. We also introduce informal discussions on parameter choices, weaknesses, related work, and open problems.

Key words: Learning with Errors, Ring Learning with Errors, Lattices, Lattice-based Cryptography, Post-quantum Cryptography.

Contents

1	Introduction	3
2	Preliminaries	5
2.1	Lattices	5
2.1.1	Key Concepts	5
2.1.2	Gaussian Distributions	6
2.1.3	Computational Problems	8
2.2	Algebraic Number Theory	10
2.2.1	Number Fields	10
2.2.2	Ring of Integers	11
2.2.3	Ideal Lattices	13
2.2.4	Efficient Computations and CRT	15
2.2.5	Ideal Lattice Problems	17
3	Learning with Errors	18
3.1	Overview	18
3.2	Regev’s Approach	20
3.2.1	The Iterative Step	22
3.2.2	The Classical Part	23
3.3	Finding an Oracle’s Center	25
3.3.1	Application to LWE	28
4	Ring-LWE	30
4.1	Overview	30
4.2	Hardness	33
4.2.1	Overview	33
4.2.2	The Iterative Step	34
4.2.3	From BDD to Ring-LWE samples	37
4.2.4	Using the Decision Oracle	38
4.3	Related Work and Variants	42
4.3.1	Weaknesses	42
4.3.2	Variants of LWE	43
4.3.3	Future Directions	44
	Bibliography	45

Chapter 1

Introduction

The Learning with Errors Problem (LWE) has been in the spotlight for theoretical computer scientists and cryptographers during the last 15 years [Reg10]. LWE has many remarkable features, but there are perhaps two major reasons why it attracts such a big interest: there exist hardness proofs based on reductions from standard lattice problems, and it allows for highly versatile cryptographic constructions. As a consequence, it is one of the main building blocks of post-quantum cryptography [Pei16a].

Learning with Errors The idea behind LWE is quite intuitive. Let \mathbb{Z}_q be the ring of integers modulo $q \geq 2$, and let $A \in \mathbb{Z}_q^{m \times n}$, $\mathbf{b} \in \mathbb{Z}_q^m$ and a secret $\mathbf{s} \in \mathbb{Z}_q^n$. Then, we can formulate a linear system of equations such as $A\mathbf{s} = \mathbf{b}$. If we are given A, \mathbf{b} , then it is easy to recover \mathbf{s} by standard methods such as Gaussian elimination. Now, suppose that we are given A and a *noisy* vector \mathbf{b} , this is, $A\mathbf{s} + \mathbf{e} = \mathbf{b}$ where $\mathbf{e} \in \mathbb{Z}_q^m$ has small coefficients (and is not known). For example, an instance with $n = 2$, $m = 3$ and $q = 13$ could be:

$$\begin{aligned} 3s_1 + 7s_2 &\approx 10 \pmod{13} \\ 2s_1 + s_2 &\approx 12 \pmod{13} \\ 11s_1 + s_2 &\approx 9 \pmod{13} \end{aligned}$$

In this case, the secret $\mathbf{s} = (11, 4)$ and the noise vector $\mathbf{e} = (1, -1, 1)$. Finding \mathbf{s} from the system of equations (samples) is no longer straightforward. Indeed, LWE is hard on both its search (find \mathbf{s} from the samples) and decision (decide whether the samples are LWE samples or uniformly random) forms, as proven by Regev in a seminal work [Reg09]. In applications, the noise is drawn from an error distribution χ which is commonly a discrete Gaussian centered at 0, such that the probability of sampling a large error (such as $|e_i| \geq q/4$) is very small. To make the problem hard, the errors must be significant. Nevertheless, they cannot be too large since they would override the LWE structure, turning the equations into random noise. The idea behind many cryptographic applications is to provide (A, \mathbf{b}) as a public key while keeping \mathbf{s} as a secret key. For instance, the original construction by Regev [Reg09] encrypts a bit by adding several equations from (A, \mathbf{b}) and hiding the bit in the last coordinate.

As mentioned before, LWE is versatile and its applications go far beyond usual public key cryptosystems. Most remarkably, LWE allows the construction of fully homomorphic encryption [Gen09] and attribute-based encryption schemes¹. However, it also presents a major drawback, which is the large sizes of both the keys and the ciphertexts required to perform encryption. Regev's cryptosystem requires key sizes of $\tilde{O}(n^2)$, and state-of-the-art schemes such as FrodoKEM [BCD⁺16] are still too expensive for everyday use in, for example, TLS or HTTPS connections.

¹See [Pei16a] for a detailed survey on lattice-based cryptography.

Ring-LWE A major question is then, how can LWE be made efficient? One answer is to use rings. Instead of noisy equations in \mathbb{Z}_q , consider an equation $as + e = b$ over a ring R . If a suitable ring is chosen (such as the ring of integers of a number field), a single equation gives more LWE samples. Besides, one can benefit from an efficient multiplication operation. We call this problem Ring Learning with Errors (Ring-LWE or simply RLWE). The main concern of using RLWE is that samples acquire an additional algebraic structure in the form of ideal lattices, induced by the ring, that might be exploited by some clever attacks. The first hardness results were established by Lyubashevsky, Peikert, and Regev [LPR10], building on the work by Regev for standard LWE [Reg09]. However, the idea of using rings and ideal lattices had appeared before in the NTRU cryptosystem [HPS98] or in early versions of Polynomial-LWE [SSTX09].

In the last decade, there have been several lines of work related to Ring-LWE. One of them is to explore other forms of algebraically structured LWE, mostly with the aim of finding a tradeoff between the security of LWE and the efficiency of RLWE. Examples are Module-LWE (MLWE), Learning with Rounding (LWR), Middle-Product LWE (MP-LWE), and Polynomial-LWE (PLWE). A second line is the construction of practical algorithms and cryptosystems, particularly for the post-quantum cryptography standardization process proposed by NIST [AASA⁺20]. Indeed, the aforementioned variants are behind final-round candidates such as Crystals-Kyber [BDK⁺18] or SABER [DKRV18]. A third line, which is the main focus of this project, is proving solid hardness results for RLWE.

This work Most of the published work on this area builds on previous results and is hard to parse for a reader new to the topic. In this report, we provide a formal introduction to Ring-LWE, presenting a proof of the strongest hardness results up to date (namely, the hardness of the decision version [PRSD17]). We focus on intuition and on drawing parallelisms between RLWE and LWE, and we also prove several non-standard results used in the hardness proof of RLWE. There are however two important results that we do not introduce in detail. These are the quantum part of Regev’s reduction, and the reductions from standard lattice problems (SIVP, GapSVP) to the Discrete Gaussian Sampling (DGS) problem. Both results appear in the original LWE reduction [Reg09] and do not need to be adapted to the ring setting.

We do not claim any new results, but aim to approach the problem in a natural and accessible way, which can be helpful for a reader that is either new to LWE, or already familiar but interested in the hardness results. The only required knowledge is basic computational complexity and undergraduate-level mathematics, although some familiarity with lattices and algebraic number theory is definitely helpful. We note that, while other LWE surveys do exist, their approaches are different to ours. Since Regev’s survey [Reg10] was published, there have been many new results, especially around Ring-LWE. Other works, such as [Pei16a], focus on cryptographic applications.

Roadmap In Chapter 2, we introduce lattices and related matters such as sampling and computational problems. We also present concepts from algebraic number theory and ideal lattices required for RLWE. In Chapter 3, we recall the original Learning with Errors problem, formalize it and present the ideas behind its hardness proof. We will not introduce RLWE formally until Chapter 4, where we study how to extend these ideas to rings, present the hardness results, and introduce some insights.

Acknowledgements I am grateful to Johan Håstad for his supervision and suggestions for improvement. I also want to thank Per Austrin, Phillip Gajland, and Maribel González-Vasco for helpful comments on the draft.

Chapter 2

Preliminaries

In this chapter, we present the most relevant concepts underlying LWE and Ring-LWE. The purpose is to provide an introduction to the topic, as well as to serve as a reference for further chapters and to clarify notation. We assume a background in linear algebra, algebraic structures, and basic computational complexity. Before introducing lattices, we introduce the statistical distance as in [MG02].

Definition 2.1 (Statistical distance). *Let X, Y be two random variables over a countable set A . The statistical distance between them is*

$$\Delta(X, Y) = \sum_{a \in A} |\Pr[X = a] - \Pr[Y = a]|.$$

If X, Y have probability density functions φ_X, φ_Y on \mathbb{R}^n ,

$$\Delta(\varphi_X, \varphi_Y) = \int_{\mathbb{R}^n} |\varphi_X(\mathbf{t}) - \varphi_Y(\mathbf{t})| d\mathbf{t}.$$

The statistical distance satisfies the triangle inequality. Besides, it cannot increase by applying any (possibly randomized) function f to the random variables, i.e., $\Delta(f(X), f(Y)) \leq \Delta(X, Y)$.

2.1 Lattices

Lattices appear not only in cryptography, but also in many areas of mathematics. For example, they play a fundamental part in algebraic number theory. Besides, lattices are used in different contexts such as for modelling crystalline structures in physics. A broader coverage of lattices can be found in [Cas59, LJ08]. In [MG02], a computational overview is presented.

2.1.1 Key Concepts

Throughout the paper, we use $\langle \cdot, \cdot \rangle$ for the inner product and $\|\cdot\|$ for the usual ℓ_2 norm. Other ℓ_p norms, including ℓ_∞ , are denoted by $\|\cdot\|_p$. For vectors $\mathbf{x} \in \mathbb{R}^n$, the standard inequality $\|\mathbf{x}\|_p \leq n^{1/p} \cdot \|\mathbf{x}\|_\infty$ holds.

Definition 2.2. *Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of \mathbb{R}^n . The lattice $\mathcal{L}(\mathcal{B})$ is the additive linear subgroup generated by the integer combinations of the vectors,*

$$\mathcal{L} = \{a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \dots + a_n\mathbf{b}_n : a_i \in \mathbb{Z}\}.$$

The basis \mathcal{B} is called the *lattice basis*. It is often represented in matrix form, $B = [\mathbf{b}_1, \dots, \mathbf{b}_n]$. The integer n is called the *dimension* of the lattice. The region enclosed by the basis vectors in the euclidean space is the *fundamental domain*, denoted by $\Lambda(\mathcal{L})$. The basis of a lattice is not unique, but there is a fundamental parameter of the lattice that does not change with the basis, called the determinant.

Definition 2.3. *Let $\mathcal{L}(B)$ be the lattice generated by B . The determinant $\det(\mathcal{L})$ of the lattice is the determinant of the lattice basis, $\det(B)$.*

The determinant can be seen as the volume of the fundamental domain, $\det(\mathcal{L}) = \text{Vol}(\Lambda(\mathcal{L}))$. A practical way of calculating the determinant of the lattice is via a Gram-Schmidt orthogonalization. If $\mathbf{b}_1, \dots, \mathbf{b}_n$ are the vectors of the basis B and $\mathbf{b}_1^*, \dots, \mathbf{b}_n^*$ are their Gram-Schmidt orthogonal vectors, then $\det(\mathcal{L}(B)) = \det(B) = \prod_{i=1}^n \|\mathbf{b}_i^*\|$.

In many applications, it is interesting to know more about short vectors and short bases of lattices. Let $\lambda_1(\mathcal{L}) = \{\inf_{\mathbf{v} \in \mathcal{L} \setminus \{0\}} \|\mathbf{v}\|\}$ be the shortest length of a nonzero vector in \mathcal{L} , that we denote by the *minimum distance*. More generally, we define $\lambda_m(\mathcal{L})$ as the smallest d such that \mathcal{L} has m independent vectors of length at most d . A theorem by Minkowski gives us a bound for λ_1 .

Theorem 2.4 (Minkowski). *In any n -dimensional lattice \mathcal{L} there exists a nonzero vector \mathbf{v} that satisfies $\|\mathbf{v}\|_\infty \leq \det(\mathcal{L})^{1/n}$. Moreover, using the standard inequality between the ℓ_∞ and ℓ_p norms, $\|\mathbf{v}\|_p \leq n^{1/p} \det(\mathcal{L})^{1/n}$.*

Unfortunately for some applications (but, as we will see, fortunately for cryptography) the proof of the theorem above does not hint any efficient method for finding short vectors. There exist lattice basis reduction algorithms such as the LLL algorithm [LLL82], which yield approximations to the shortest vector. The first vector \mathbf{v}_1 of a *LLL-reduced basis* of \mathcal{L} is guaranteed to satisfy $\|\mathbf{v}_1\| \leq 2^{(n-1)/2} \lambda_1(\mathcal{L})$. Another important concept that arises frequently in lattice-related constructions is the dual lattice.

Definition 2.5. *The dual lattice \mathcal{L}^* of \mathcal{L} is the lattice*

$$\mathcal{L}^* = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbf{v} \rangle \in \mathbb{Z}, \forall \mathbf{v} \in \mathcal{L}\}.$$

If B is a basis of \mathcal{L} , then $(B^{-1})^t$ is a basis of \mathcal{L}^* . Hence, $\det(\mathcal{L}^*) = \det(\mathcal{L})^{-1}$. Given a dual lattice \mathcal{L}^* , we say that \mathcal{L} is its primal lattice. The following “transference theorem” is a consequence of the primal-dual correspondence.

Lemma 2.6 (Duality [Ban93]). *For any n -dimensional lattice \mathcal{L} , $1 \leq \lambda_1(\mathcal{L}) \cdot \lambda_n(\mathcal{L}^*) \leq n$.*

2.1.2 Gaussian Distributions

A central tool for working with Learning with Errors are discrete and continuous Gaussian distributions, as well as sampling from them. Let $r > 0$ be a width parameter and $\rho_r : \mathbb{R}^n \rightarrow (0, 1]$ be a Gaussian function, given by $\rho_r(\mathbf{x}) = \exp(-\pi\|\mathbf{x}\|^2/r^2)$. By normalizing it, we obtain the n -dimensional (spherical) Gaussian distribution,

$$D_r = \frac{\rho_r}{\int_{\mathbf{x} \in \mathbb{R}^n} \rho_r(\mathbf{x}) d\mathbf{x}} = \rho_r r^{-n}.$$

A sample from the n -dimensional Gaussian distribution can be obtained from taking n independent samples from the 1-dimensional case. This fact allows for an easy extension to

elliptical Gaussian distributions, where the width is no longer a scalar but a vector¹. Let $\mathbf{r} = (r_1, \dots, r_n) \in (\mathbb{R}^+)^n$ and let $\{\mathbf{h}_i\}_i$ be an orthonormal basis of \mathbb{R}^n . Then, a sample from $D_{\mathbf{r}}$ in the basis $\{\mathbf{h}_i\}_i$ is given by $\sum_{i=1}^n x_i \mathbf{h}_i$, where the $x_i \in \mathbb{R}$ are sampled from the 1-dimensional D_{r_i} . The Gaussian function in this case is given by $\rho_{\mathbf{r}}(\mathbf{x}) = \exp(-\pi \sum_{i=1}^n |x_i/r_i|^2)$.

Lemma 2.7 (Statistical distance [Reg09]). *For $0 < \alpha < \beta$, the statistical distance between D_{α} and D_{β} is $\Delta(D_{\alpha}, D_{\beta}) \leq 10(\beta/\alpha - 1)$.*

We can also define a *discrete Gaussian* probability distribution $D_{\mathcal{L}, r}$ over a lattice \mathcal{L} , given by

$$D_{\mathcal{L}, r}(\mathbf{v}) = \frac{\rho_{\mathbf{r}}(\mathbf{v})}{\rho_{\mathbf{r}}(\mathcal{L})}, \quad \forall \mathbf{v} \in \mathcal{L}.$$

Note that a discrete Gaussian can be defined over any discrete set of points, and not only over a lattice. The *smoothing parameter* is a lattice constant that provides the width beyond which a discrete Gaussian behaves like a continuous Gaussian. The definition is not very intuitive, but the lemma that follows captures the idea behind it. We introduce it for the spherical case.

Definition 2.8 (Smoothing parameter). *Let \mathcal{L} be a lattice and let $r, \epsilon > 0$. The smoothing parameter $\eta_{\epsilon}(\mathcal{L})$ is the smallest r such that $\rho_{1/r}(\mathcal{L}^* \setminus \{0\}) \leq \epsilon$.*

Lemma 2.9. *For a lattice \mathcal{L} , a vector $\mathbf{c} \in \mathbb{R}^n$, $\epsilon \geq 0$ and $r \geq \eta_{\epsilon}(\mathcal{L})$, the statistical distance between $D_r + \mathbf{c} \bmod \mathcal{L}$ and the uniform distribution over \mathbb{R}^n/\mathcal{L} is at most $\epsilon/2$. Besides, $\rho_r(\mathcal{L} + \mathbf{c}) \in \left[\frac{1-\epsilon}{1+\epsilon}, 1\right] \cdot \rho_r(\mathcal{L})$.*

The smoothing parameter is one of the most important concepts that arise in lattices and LWE. Let us try to develop the intuition behind Definition 2.8. For a large r , the distribution $\rho_{1/r}$ will be very narrow, the zero vector being by far the most probable one. When r decreases, the distribution will become wider and the probability of sampling from $\rho_{1/r}(\mathcal{L}^* \setminus \{0\})$ will no longer be negligible. In the primal lattice \mathcal{L} , this means that ρ_r will become narrower and the discrete Gaussian structure will be more significant.

The sparsity of the lattice also influences the smoothing parameter. If the points in \mathcal{L} are distant from each other, \mathcal{L}^* will be dense and r is required to be large for $\rho_{1/r}$ to sample only the zero vector. Indeed, in sparse lattices, we will need our discrete Gaussian to be wider in order to hide the discrete structure. On a different note, it is not hard to see that $\eta_{\epsilon}(\mathcal{L})$ is a continuous and decreasing function of ϵ .

An equivalent way of describing the smoothing parameter is as the smallest continuous Gaussian noise required to hide a lattice structure, as seen in the following lemma.

Lemma 2.10 (Sum of distributions [Reg09]). *Let \mathcal{L} be a lattice, $r, s > 0$ be reals and $\mathbf{u} \in \mathbb{R}^n$ any vector. Assume that $1/\sqrt{1/r^2 + 1/s^2} \geq \eta_{\epsilon}(\mathcal{L})$ for some $\epsilon < \frac{1}{2}$. Consider the continuous distribution Y on \mathbb{R}^n obtained by sampling from $D_{\mathcal{L}+\mathbf{u}, r}$ and then adding an element drawn independently from D_s . Then the statistical distance between Y and $D_{\sqrt{r^2+s^2}}$ is at most 4ϵ .*

As a rule of thumb, in LWE and Ring-LWE it is necessary to work with distributions whose width is above the smoothing parameter, to guarantee that operations such as the sum of distributions behave adequately. We can also define an elliptical version of the smoothing parameter (that we name smoothing condition), which will be used only for some parts in Chapter 4. Lemmas 2.9 and 2.10 extend easily to the elliptical case [PRSD17].

¹Indeed, elliptical Gaussians are a particular case of multivariate Gaussian distributions characterized by positive definite symmetric matrices. These distributions are common in the Ring-LWE setting [RSW18], but will not be needed in this work.

Definition 2.11 (Smoothing condition [PRSD17]). *Let \mathcal{L} be a lattice and let $\mathbf{r} \in (\mathbb{R}^+)^n$. We write $\mathbf{r} \geq \eta_\epsilon(\mathcal{L})$ if $\rho_{1/\mathbf{r}}(\mathcal{L}^* \setminus \{0\}) \leq \epsilon$, where $1/\mathbf{r} = (1/r_1, \dots, 1/r_n)$.*

Lemma 2.12 ([MR07, Reg09, PRSD17]). *Let $c \geq 1$. For any n -dimensional lattice \mathcal{L} and $\epsilon = \exp(-c^2 n)$, we have that*

$$\sqrt{\pi^{-1} \ln(1/\epsilon)} \leq \lambda_1(\mathcal{L}^*) \cdot \eta_\epsilon(\mathcal{L}) \leq c\sqrt{n}.$$

In particular, this implies that for $\epsilon = 2^{-O(n)}$, we have $\eta_\epsilon(\mathcal{L}) \geq O(\sqrt{n})/\lambda_1(\mathcal{L}^*)$. Unless explicitly stated, η_ϵ will refer to the smoothing parameter in the (spherical) standard case.

A critical part of lattice-based cryptography resides in the ability to sample from discrete Gaussian distributions. Sampling from $D_{\mathcal{L}, \mathbf{r}}$ is only feasible when \mathbf{r} is very large (unless a very short basis is known). Otherwise, many computational problems in lattices would be easy, as we will see in Section 2.1.3.

Lemma 2.13 (Long vectors are unlikely [Ban93, LPR10]). *For any n -dimensional lattice \mathcal{L} and a radius $r > 0$, a sample s from $D_{\mathcal{L}, r}$ satisfies $\|s\| \leq r\sqrt{n}$, except with probability at most 2^{-2n} .*

There do exist efficient algorithms to sample from wide discrete Gaussian distributions. For this work, it suffices to present the original bootstrapping method presented in [Reg09], which is summarized in the following lemma.

Lemma 2.14 (Bootstrapping [Reg09]). *Given any n -dimensional lattice \mathcal{L} and a vector \mathbf{r} such that $r_i \geq 2^{2n} \lambda_n(\mathcal{L})$ for every coordinate i , there exists an efficient algorithm that outputs an independent sample from a distribution Y such that the statistical distance $\Delta(D_{\mathcal{L}, \mathbf{r}}, Y) \leq 2^{-\Omega(n)}$.*

The bootstrapping algorithm consists in finding an LLL-reduced basis B of \mathcal{L} (whose vectors have length at most $2^n \lambda_n(\mathcal{L})$). Then, we can sample a vector \mathbf{y} from D_r and output $\mathbf{y} - (\mathbf{y} \bmod B) \in \mathcal{L}$, which is proven to be exponentially close to $D_{\mathcal{L}, \mathbf{r}}$ (recall that the coordinates of \mathbf{r} are very large). The proof in [Reg09] is presented only for spherical Gaussians, but it holds when D_r is replaced by the continuous elliptical distribution $D_{\mathbf{r}}$. In applications, sampling can be done in more efficient ways, especially when a shorter basis is provided [Pei10].

2.1.3 Computational Problems

There are several hard computational lattice problems that are important for LWE, all of them related to finding short vectors (or equivalently, short lattice bases).

Definition 2.15 (Shortest Vector Problem, SVP). *Given a basis $B \in \mathbb{R}^{n \times n}$ of a lattice $\mathcal{L}(B)$, find a nonzero lattice vector $\mathbf{x} \in \mathcal{L}$ such that $\|\mathbf{x}\| \leq \|\mathbf{v}\|$ for any $\mathbf{v} \in \mathcal{L} \setminus \{0\}$.*

Definition 2.16 (Closest Vector Problem, CVP). *Given a basis $B \in \mathbb{R}^{n \times n}$ of a lattice $\mathcal{L}(B)$, and a vector $\mathbf{y} \in \mathbb{R}^n$, find the closest lattice vector to \mathbf{y} , i.e., find $\mathbf{x} \in \mathcal{L}$ such that $\|\mathbf{y} - \mathbf{x}\| \leq \|\mathbf{y} - \mathbf{v}\|$ for any $\mathbf{v} \in \mathcal{L}$.*

The previous formulation is the *search* version of the SVP and CVP problems. It is important to distinguish them from their respective decision versions, which is easier (the hardness of the decision problems implies the hardness of the search problem [MG02]). The SVP (resp. CVP) decision problem can be formulated as, given a constant $c > 0$, decide whether there is a nonzero vector $\mathbf{x} \in \mathcal{L}$ such that $\|\mathbf{x}\| \leq c$ (resp. $\|\mathbf{x} - \mathbf{v}\| \leq c$).

SVP and CVP are often found in their approximation form, especially in cryptography. These are parametrized by an approximation factor $\gamma \geq 1$ which is generally a function of the lattice dimension, $\gamma = \gamma(n)$ [Pei16a]. These approximation versions are:

Definition 2.17 (Approximate SVP, SVP_γ). *Given a basis $B \in \mathbb{R}^{n \times n}$ of a lattice $\mathcal{L}(B)$, find a nonzero lattice vector $\mathbf{x} \in \mathcal{L}$ such that $\|\mathbf{x}\| \leq \gamma \|\mathbf{v}\|$ for any other $\mathbf{v} \in \mathcal{L} \setminus \{0\}$.*

The definition for CVP_γ is analogous. We can formulate *promise* versions of SVP_γ and CVP_γ , denoted by $\text{GapSVP}_\gamma, \text{GapCVP}_\gamma$.

Definition 2.18 (Promise $\text{SVP}_\gamma, \text{GapSVP}_\gamma$). *An instance of GapSVP_γ is a pair (B, c) where $B \in \mathbb{R}^{n \times n}$ is a lattice basis and c is a positive real number, such that*

- (B, c) is a YES instance if $\lambda_1(\mathcal{L}(B)) \leq c$.
- (B, c) is a NO instance if $\lambda_1(\mathcal{L}(B)) > \gamma c$.

The Shortest Independent Vectors problem can be seen as a generalization of SVP for a set of lattice generators. We introduce it in its approximation form.

Definition 2.19 (Shortest Independent Vectors Problem, SIVP_γ). *Given a basis $B \in \mathbb{R}^{n \times n}$ of a lattice $\mathcal{L}(B)$, find n linearly independent short vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ such that $\max_i \|\mathbf{v}_i\| \leq \gamma \lambda_n(\mathcal{L})$.*

The CVP problem is NP-complete, and in fact it is hard to approximate within quasi-polynomial factors [DKS98]. However, so far, no cryptosystem has been proven to be secure based on a CVP-related problem [Pei16a]. The results for SVP are somewhat weaker. Khot [Kho05] proves that SVP is NP-hard for any constant factor under randomized reductions. We note that an applicable hardness result for the LWE problem (which is based on the GapSVP_γ and SIVP_γ problems) would require polynomial factors.

We now list three less common lattice problems that arise when studying LWE and RLWE [Reg09, PRSD17].

Definition 2.20 (Bounded Distance Decoding, $\text{BDD}_{\mathcal{L},d}$). *Given a lattice \mathcal{L} , a distance bound $d < \lambda_1(\mathcal{L})/2$, and \mathbf{y} of the form $\mathbf{y} = \mathbf{x} + \mathbf{e}$ for $\mathbf{x} \in \mathcal{L}$ and $\|\mathbf{e}\| \leq d$, find the lattice vector \mathbf{x} (or, equivalently, find \mathbf{e}).*

Definition 2.21 (Discrete Gaussian Sampling, DGS_φ). *Given a lattice \mathcal{L} , a function $\varphi : \mathcal{L} \rightarrow \mathbb{R}^+$, and a parameter $r \leq \varphi(\mathcal{L})$, output an independent sample from a distribution which is within statistical distance of $D_{\mathcal{L},r}$.*

Definition 2.22 (Gaussian Decoding Problem, $\text{GDP}_{\mathcal{L},g}$). *Given a lattice \mathcal{L} , a Gaussian parameter $g > 0$, and \mathbf{y} of the form $\mathbf{y} = \mathbf{x} + \mathbf{e}$ for $\mathbf{x} \in \mathcal{L}$ where $\mathbf{e} \in \mathbb{R}^n$ is drawn from D_g , find the lattice vector \mathbf{x} (or, equivalently, find \mathbf{e}).*

The $\text{BDD}_{\mathcal{L},d}$ and GapCVP_γ problems can be seen as duals of each other. $\text{BDD}_{\mathcal{L},d}$ is easier to solve when we are promised a point which is close to the lattice, i.e., when we have a small d . On the other hand, the GapCVP_γ problem becomes easier when γ increases. In fact, for $\text{BDD}_{\mathcal{L},d}$ we have the following result, which can be seen as a “dual” form of the the LLL algorithm.

Lemma 2.23 (Babai Nearest Plane [Bab86]). *There exists a polynomial time algorithm that solves $\text{BDD}_{\mathcal{L},d}$ for $d \leq 2^{-n/2} \lambda_1(\mathcal{L})$.*

The DGS_φ problem is closely related to problems related to finding short vectors. There exist polynomial-time reductions from the standard problems SIVP_γ and GapSVP_γ to DGS_φ , as introduced in [Reg09]. Roughly, this is because if we solve DGS_φ for a small r , we will be able to sample from a narrow discrete Gaussian distribution on \mathcal{L} , which means that we will

obtain short lattice vectors. As the sampling is random, the probability of finding a basis (of independent vectors) among such short vectors is very high, if a sufficient number of them is sampled. The reduction from GapSVP_γ follows a similar high-level idea, but requires further work.

The last of the listed problems, $\text{GDP}_{\mathcal{L},g}$, is a variant of $\text{BDD}_{\mathcal{L},d}$ in which the deviation from the lattice is Gaussian. In this respect, we introduce a self-reducibility result for the $\text{GDP}_{\mathcal{L},g}$ problem, without proving it. The result is not a fundamental part of any of the results we present in this work, but it is used to achieve tighter bounds in the reductions in Chapter 4.

Proposition 2.24 (Self-reducibility of GDP [PRSD17]). *Assume we have an oracle that solves $\text{GDP}_{\mathcal{L},g}$ with non-negligible probability over the choice of the vector \mathbf{y} . Then, we can efficiently solve $\text{GDP}_{\mathcal{L},\delta g}$ with overwhelming probability, where $\delta = \delta(n)$ is an $o(1)$ function.*

2.2 Algebraic Number Theory

The original Learning with Errors problem does not require any abstract algebra or number theory, but rather results on lattices and sampling. Therefore, it is possible to read Chapter 3 without any algebraic background. However, this is required as soon as the rings are introduced.

In this section, we review the concepts and results needed for studying the Ring Learning with Errors problem. Most of the material (but for a few exceptions, that we prove) is standard and can be found in any introductory book or notes in algebraic number theory. Our main references are [Mil08, Rib13].

2.2.1 Number Fields

We start with the definitions of algebraic number and algebraic integer, to later introduce concepts such as embedding, trace, and norm.

Definition 2.25. *We say that $\alpha \in \mathbb{C}$ is an algebraic number if there exists a polynomial $f(x) \in \mathbb{Q}[x]$ such that $f(\alpha) = 0$.*

The *minimal polynomial* of an algebraic number α is the monic irreducible polynomial that has α as a root. This polynomial is unique and denoted by $f_\alpha(x)$. Indeed, for any other $g(x) \in \mathbb{Q}[x]$ such that $g(\alpha) = 0$, we have that f_α divides g (denoted by $f_\alpha \mid g$).

If $\deg(f_\alpha) = n$, we say that the algebraic number α has degree n . The roots $\alpha^{(1)}, \dots, \alpha^{(n)}$ of its minimal polynomial $f_\alpha(x)$ are called *conjugates* of α .

Definition 2.26. *An algebraic number $\alpha \in \mathbb{C}$ is an algebraic integer if its minimal polynomial has integer coefficients.*

A *number field* K is an extension of the field of rational numbers \mathbb{Q} , that can be obtained by adjoining an algebraic number to it ($K = \mathbb{Q}(\alpha)$). All number fields can be generated by a single algebraic number (this fact is a consequence of the Primitive Element Theorem). The *degree* n of a number field is the degree of the extension $n = [K : \mathbb{Q}]$, which equals the degree of α . A number field can be seen as an n -dimensional vector space over \mathbb{Q} with basis $\{1, \alpha, \dots, \alpha^{n-1}\}$, called the *power basis* of K . This basis is not unique, and every element $\beta \in K$ has degree d divisible by n .

There is a natural field isomorphism $\varphi : \mathbb{Q}[x]/(f_\alpha(x)) \rightarrow K = \mathbb{Q}(\alpha)$, given by $\varphi(f) \mapsto f(\alpha)$. This means that the elements of a number field can be represented uniquely by polynomials;

in fact, every number field can be seen as a field of polynomials of the form $\mathbb{Q}[x]/(f_\alpha(x))$. Hence, one can think of field elements as polynomials of degree smaller than n without loss of generality. To see why this is true, consider the extension $\tilde{\varphi}$ of this map $\tilde{\varphi} : \mathbb{Q}[x] \rightarrow K$; its kernel is the set of polynomials in $\mathbb{Q}[x]$ such that $f(\alpha) = 0$. We know that $f_\alpha \mid g$ for any other $g(x) \in \mathbb{Q}[x]$ that vanishes on α . Hence, $g(x) \in (f(x))$ (i.e., the ideal generated by $f(x)$), so $\ker(\tilde{\varphi}) = (f(x))$, and φ is an isomorphism by the first isomorphism theorem.

Example 2.27. If we consider $\alpha = \sqrt{3}$, we have that α is a root of $g(x) = x^3 + x^2 - 3x - 3$, but $g(x)$ is not irreducible. If we take $f(x) = x^2 - 3$, then f is irreducible on $\mathbb{Q}[x]$ and has α as a root (notice that $f \mid g$). Therefore, f is the minimal polynomial of α , and since it has integer coefficients, then α is an algebraic integer. Adjoining α to \mathbb{Q} gives a number field of degree 2, which is an n -dimensional vector space over \mathbb{Q} whose basis is $\{1, \sqrt{3}\}$.

In a number field $K = \mathbb{Q}(\alpha)$ of degree n , there exist n distinct *embeddings* $\sigma_i : K \rightarrow \mathbb{C}$ that map α to each of its conjugates. They are given by $\sigma_i(\alpha^{(1)}) = \alpha^{(i)}$. These embeddings are \mathbb{Q} -isomorphisms, since $\sigma_i(x) = x$ for any $x \in \mathbb{Q}$. If the conjugate $\alpha^{(i)} \in \mathbb{R}$, we say that σ_i is a *real embedding*. Otherwise, it is a *complex embedding*. As the $\alpha^{(i)}$ are roots of a polynomial in $\mathbb{Q}[x]$, the complex roots always come in pairs, and so do the complex embeddings. Hence, we can denote the number of real embeddings by s_1 and the number of pairs of complex embeddings by s_2 , having $n = s_1 + 2s_2$. For convenience, we can also define an ordering of the embeddings, σ_i for $i = 1, \dots, s_1$ are the real embeddings, and $\sigma_{i+s_1+s_2} = \bar{\sigma}_{i+s_1}$ for $i = 1, \dots, s_2$ are the complex embeddings. This leads to the following definition.

Definition 2.28. *The canonical embedding is a map $\sigma : K \rightarrow \mathbb{R}^{s_1} \times \mathbb{C}^{2s_2}$ defined as $\sigma(x) = (\sigma_1(x), \dots, \sigma_n(x))$.*

In order to make σ a ring homomorphism, the multiplication of embedded elements is coordinate-wise. We can also define a (field) trace and norm for the elements of the field.

Definition 2.29. *The field norm of $\alpha \in K$, denoted by $N_{K/\mathbb{Q}}(\alpha)$ or simply by $N(\alpha)$, is the product of the conjugates of α , $N(\alpha) = \alpha^{(1)} \cdots \alpha^{(n)}$. Equivalently, $N(\alpha) = \prod_{i=1}^n \sigma_i(\alpha)$.*

Definition 2.30. *The field trace of $\alpha \in K$, denoted by $Tr_{K/\mathbb{Q}}(\alpha)$ or simply by $Tr(\alpha)$, is the sum of the conjugates of α , $Tr(\alpha) = \alpha^{(1)} + \cdots + \alpha^{(n)}$. Equivalently, $Tr(\alpha) = \sum_{i=1}^n \sigma_i(\alpha)$.*

Both the trace and the norm of an algebraic number are rational numbers. Besides, if α is an algebraic integer, they are integers (notice that they are exactly the coefficients of degree 0 and $n - 1$ of the minimal polynomial $f_\alpha(x) = (x - \alpha^{(1)}) \cdots (x - \alpha^{(n)})$). The trace is additive and the norm is multiplicative: $Tr(\alpha + \beta) = Tr(\alpha) + Tr(\beta)$ and $N(\alpha\beta) = N(\alpha)N(\beta)$ for any $\alpha, \beta \in K$.

The field norm should not be confused with the norm of an embedded element in $\mathbb{R}^{s_1} \times \mathbb{C}^{2s_2}$, (or with the norm of the lattice point associated to an algebraic integer, as we will see later). Norms of embedded elements are Euclidean norms and they are not multiplicative.

2.2.2 Ring of Integers

The sum, subtraction and product of any two algebraic integers yields another algebraic integer. Therefore, the algebraic integers of a number field K (i.e., the algebraic integers that are in K) form a ring \mathcal{O}_K , called *ring of integers* (of K). The simplest example of a ring of integers is $\mathbb{Z} = \mathcal{O}_{\mathbb{Q}}$. Besides, the ring of algebraic integers is a \mathbb{Z} -module of range n (the degree of the field), since each element can be expressed as an integer linear combination of some basis $\mathcal{B} = \{b_1, \dots, b_n\} \subset \mathcal{O}_K$. The set \mathcal{B} is called an *integral basis* of \mathcal{O}_K and is also a basis for K as a \mathbb{Q} -vector space.

It is not hard to see that the *units* (invertible elements) of the ring are those elements $u \in \mathcal{O}_K$ such that $N(u) = \pm 1$. Every element of \mathcal{O}_K admits a factorization into irreducible elements and units, but it is not necessarily unique as in \mathbb{Z} . To recover uniqueness, we have to look at how ideals decompose.

Definition 2.31. *An integral ideal $\mathcal{I} \subset \mathcal{O}_K$ is an additive subgroup of \mathcal{O}_K which is closed under multiplication in \mathcal{O}_K , i.e., if $a \in \mathcal{I}$ then $ab \in \mathcal{I}$ for any $b \in \mathcal{O}_K$.*

The integral ideals of a number field have many useful properties. Any integral ideal is finitely generated by elements in \mathcal{O}_K , and we say that it is principal whenever $\mathcal{I} = (w)$, i.e., it is generated by a single element $w \in \mathcal{O}_K$. Moreover, integral ideals also admit a basis $\{u_1, \dots, u_n\} \subset \mathcal{O}_K$ and are \mathbb{Z} -modules of \mathcal{O}_K , where each of the u_i can be respectively written in a basis of \mathcal{O}_K . For any integral ideals \mathcal{I}, \mathcal{J} , both their sum $\mathcal{I} + \mathcal{J} = \{i + j : i \in \mathcal{I}, j \in \mathcal{J}\}$ and the product $\mathcal{I}\mathcal{J} = \{i \cdot j : i \in \mathcal{I}, j \in \mathcal{J}\}$ are also integral ideals. Notice that the $\mathcal{I} + \mathcal{J}$ contains \mathcal{I} and \mathcal{J} , while their product is contained in them. Integral ideals can be uniquely factored into a product of powers of prime ideals.

Definition 2.32. *An integral ideal $\mathfrak{p} \subset \mathcal{O}_K$ is prime if whenever $ab \in \mathfrak{p}$ for $a, b \in \mathcal{O}_K$, then either $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$.*

We say that two ideals \mathcal{I}, \mathcal{J} are coprime when $\mathcal{I} + \mathcal{J} = 1$, or equivalently if they do not share any common prime factor. On rings of integers, every prime ideal is maximal, meaning that for any prime ideal \mathfrak{p} , the quotient ring $\mathcal{O}_K/\mathfrak{p}$ is a field.

Proposition 2.33. *Every integral ideal $\mathcal{I} \subset \mathcal{O}_K$ can be written uniquely as a product of the form $\mathcal{I} = \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r}$ where the \mathfrak{p}_i are prime ideals and the a_i are positive integers.*

Proposition 2.34. *For every integral ideal $\mathcal{I} \subset \mathcal{O}_K$, there exists another ideal $\mathcal{I}' \subset \mathcal{O}_K$ such that $\mathcal{I}\mathcal{I}' = (w)$ is principal.*

An important property of ideals is their norm, which is a generalization of the field norm defined in Definition 2.29.

Definition 2.35. *The ideal norm $N(\mathcal{I})$ of an integral ideal $\mathcal{I} \subset \mathcal{O}_K$ is its index as an ideal of \mathcal{O}_K , $N(\mathcal{I}) = |\mathcal{O}_K/\mathcal{I}|$.*

The norm is multiplicative, such that $N(\mathcal{I}\mathcal{J}) = N(\mathcal{I}) \cdot N(\mathcal{J})$. Besides, we have that if $\mathcal{I} = (w)$ is principal, then $N(\mathcal{I}) = N_{K/\mathbb{Q}}(w)$ (hence the generalization of the field norm). Notice that this also means that for any $x \in \mathcal{I}$, then $N(\mathcal{I})$ divides $N(x)$ as $(x) \subset \mathcal{I}$. Despite all these nice properties, the set of ideals does not have a group structure, i.e., integral ideals are not invertible. To recover it, we need fractional ideals.

Definition 2.36. *A fractional ideal (also called \mathcal{O}_K -ideal) is an \mathcal{O}_K -module $\mathcal{I} \subset K$ such that $d\mathcal{I}$ is an integral ideal for some $d \in \mathcal{O}_K$. Equivalently (for number fields), a fractional ideal is a finitely generated \mathcal{O}_K -submodule of K .*

Notice that this definition includes integral ideals. The set of fractional ideals is indeed a multiplicative group; it is easy to check that the product is closed and that the ring \mathcal{O}_K is a unit element. For an integral ideal \mathcal{I} , its inverse $\mathcal{I}^{-1} = w^{-1}\mathcal{I}'$ where \mathcal{I}' and w are defined as in Proposition 2.34. For a fractional ideal \mathcal{I} , then $d\mathcal{I}$ is integral for some $d \in \mathcal{O}_K$ and $\mathcal{I}^{-1} = d^{-1}(d\mathcal{I})^{-1}$.

Both the ideal norm and the unique factorization can be extended to fractional ideals. For a fractional ideal \mathcal{I} such that $d\mathcal{I} \subset \mathcal{O}_K$, $N(\mathcal{I}) = N(d\mathcal{I}) \cdot N(d)^{-1}$. For the unique factorization, we have that $\mathcal{I} = (d)^{-1}(d\mathcal{I})$. Hence, if $d\mathcal{I} = \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r}$ and $(d) = \mathfrak{q}_1^{b_1} \cdots \mathfrak{q}_s^{b_s}$, then $\mathcal{I} = \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r} \mathfrak{q}_1^{-b_1} \cdots \mathfrak{q}_s^{-b_s}$. Besides, fractional ideals also admit a basis $\{u_1, \dots, u_n\} \subset K$.

We remark that fractional ideals are *not* ideals of the ring of algebraic integers, in spite of their name.

2.2.3 Ideal Lattices

Recall the canonical embedding σ from Definition 2.28. The image of K over σ actually lives in a particular space $H \subset \mathbb{C}^n$ defined as

$$H = \{(x_1, \dots, x_n) \in \mathbb{R}^{s_1} \times \mathbb{C}^{2s_2} : x_{s_1+s_2+j} = \overline{x_{s_1+j}}, \forall j \in [s_2]\} \subset \mathbb{R}^{s_1} \times \mathbb{C}^{2s_2}.$$

The space H is isomorphic to \mathbb{R}^n via an orthonormal basis $\{\mathbf{h}_1, \dots, \mathbf{h}_n\}$, that we can define from the canonical basis $\{\mathbf{e}_i\}_i \subset \mathbb{C}^n$ as follows: for $1 \leq i \leq s_1$, let $\mathbf{h}_i = \mathbf{e}_i$, and for $s_1 < i \leq s_1 + s_2$, take $\mathbf{h}_i = \frac{1}{\sqrt{2}}(\mathbf{e}_i + \mathbf{e}_{i+s_2})$ and $\mathbf{h}_{i+s_2} = \frac{\sqrt{-1}}{\sqrt{2}}(\mathbf{e}_i - \mathbf{e}_{i+s_2})$. The norm of an element of H is the norm induced by \mathbb{C}^n ; in the case of the ℓ_2 norm, it corresponds exactly with the norm on \mathbb{R}^n given by the isomorphism. Generally speaking, the ℓ_p norm of a field element $x \in K$ under the embedding is

$$\|x\|_p = \|\sigma(x)\|_p = \left(\sum_{i=1}^n |\sigma_i(x)|^p \right)^{1/p},$$

and the ℓ_∞ norm is $\|x\|_\infty = \max_i |\sigma_i(x)|$ (again, we remark that this is different from the field norm in Definition 2.29). We also have that for $x, y \in K$ and any ℓ_p norm, $\|xy\|_p \leq \|x\|_\infty \cdot \|y\|_p$.

H is a larger space than K , since K is a \mathbb{Q} -vector space. However, there is an isomorphism between the spaces $H \cong K_{\mathbb{R}}$, where $K_{\mathbb{R}} = K \otimes_{\mathbb{Q}} \mathbb{R}$ is a tensor product space (intuitively, one can think of this tensor product as extending the \mathbb{Q} -vector space structure of K to an \mathbb{R} -vector space, by replacing rational coefficients by reals). This property allows the extension of the results in Section 2.1.2 to Gaussian distributions over number fields. Following the notation in [PRSD17], let $G = \{\mathbf{r} \in (\mathbb{R}^+)^n : r_{i+s_1+s_2} = r_{i+s_1}, i = 1, \dots, s_2\}$ be a set of possible radial vectors, and let $D_{\mathbf{r}}$ be defined by the basis $\{\mathbf{h}_i\}_i$. For any $\mathbf{r} \in G$, $D_{\mathbf{r}}$ is a distribution over $H \cong K_{\mathbb{R}}$. For any element $x \in K_{\mathbb{R}}$, the distribution $x \cdot D_{\mathbf{r}}$ is $D_{\mathbf{r}'}$, where the components $r'_i = r_i \cdot |\sigma_i(x)|$, using the constraint on the choice of \mathbf{r} . When working with (elliptical) Gaussian distributions over number fields, we will always consider this setting.

In H , one can embed not only field elements, but also (fractional) ideals.

Definition 2.37 (Ideal lattice). *Let \mathcal{I} be a fractional ideal with basis $\{u_1, \dots, u_n\}$. Its image under the canonical embedding is a lattice $\sigma(\mathcal{I}) \subset H$ given by the basis $\{\sigma(u_1), \dots, \sigma(u_n)\}$. Lattices of this form are called ideal lattices.*

For readability, we will often denote $\sigma(\mathcal{I})$ by \mathcal{I} . Ideal lattices are discrete additive subgroups of H , and we can extend the discrete Gaussian distributions from Section 2.1.2 to them.

Example 2.38. We illustrate the concepts of ideals, ideal basis, the canonical embedding and ideal lattices with a toy example. Let $K = \mathbb{Q}(i)$ be the fourth cyclotomic field, generated by the polynomial $(x^2 + 1)$. K has degree 2 and the conjugates of i are i and $-i$. We have that $\mathcal{O}_K = \mathbb{Z}[i]$, hence a basis of \mathcal{O}_K is $\{1, i\}$ and any algebraic integer can be written as $a + bi$ for $a, b \in \mathbb{Z}$. Alternatively, we can look at K, \mathcal{O}_K as polynomial fields: $K = \mathbb{Q}[x]/(x^2 + 1)$ and $\mathcal{O}_K = \mathbb{Z}[x]/(x^2 + 1)$.

We want to study how an ideal $\mathcal{I} \subset \mathcal{O}_K$ embeds into an ideal lattice in H . For this purpose, let us choose the principal ideal $\mathcal{I} = (-2 + i)$. The first step is to write this ideal as a \mathbb{Z} -module on the basis of \mathcal{O}_K . Any element $\alpha \in \mathcal{I}$ is a multiple of $(-2 + i)$, therefore it is of the form $(a + bi)(-2 + i) = a(-2 + i) + b(-1 - 2i)$ for $a, b \in \mathbb{Z}$. The elements $-2 + i$ and $-1 - 2i$ form a basis of \mathcal{I} .

Finally, we need to map our ideal into H via the canonical embedding. Notice that $\mathbf{v}_1 = \sigma(1) = (1, 1)$ and $\mathbf{v}_2 = \sigma(i) = (i, -i)$, therefore $\sigma(\mathcal{O}_K) \subset H$ is the lattice generated by $\mathbf{v}_1, \mathbf{v}_2$. Notice that both of our embeddings are complex, hence $H \subset \mathbb{C}^2$. The lattice $\sigma(\mathcal{I})$ is a sublattice of $\sigma(\mathcal{O}_K)$ which is determined by the canonical embedding of the basis of \mathcal{I} , i.e., by $\sigma(-2 + i) = (-2 + i, 2 + i)$ and $\sigma(-1 - 2i) = (-1 - 2i, -1 + 2i)$. Thus, we have

$$\sigma(\mathcal{I}) = \{a(-2 + i, 2 + i) + b(-1 - 2i, -1 + 2i) : a, b \in \mathbb{Z}\} \subset \sigma(\mathcal{O}_K) \subset H.$$

If we express the basis of $\sigma(\mathcal{I})$ with respect to the basis of $\sigma(\mathcal{O}_K)$, we obtain that $\sigma(\mathcal{I})$ is generated by the matrix basis $B = \begin{bmatrix} -2 & 1 \\ -1 & -2 \end{bmatrix}$. Since $\det(B) = 5$, the fundamental volume $\text{Vol}(\sigma(\mathcal{I})) = 5 \text{Vol}(\sigma(\mathcal{O}_K))$; one out of five lattice points of $\sigma(\mathcal{O}_K)$ is in $\sigma(\mathcal{I})$, as seen in Figure 2.1.

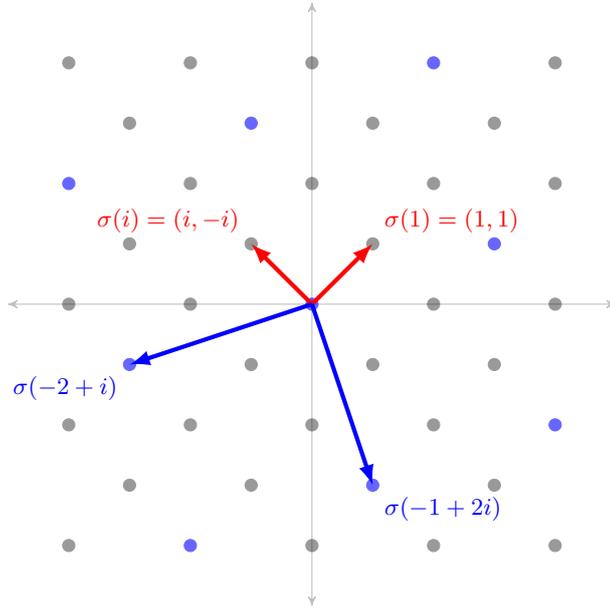


Figure 2.1: Ideal lattice $\sigma(\mathcal{I})$, where $\mathcal{I} = (-2 + i)$ and $\mathcal{O}_K = \mathbb{Z}[i]$.

The *discriminant* Δ_K of a number field K is the square of the fundamental volume of the ideal lattice generated by the ring of integers \mathcal{O}_K . Intuitively, it measures the sparsity of the algebraic integers of the ring. Alternatively, we have that $\Delta_K = |\det(\text{Tr}(b_i \cdot b_j))|$ where the b_i form an integral basis of \mathcal{O}_K . Hence, the fundamental volume of an ideal lattice $\mathcal{I} \subset \mathcal{O}_K$ is $\det(\mathcal{I}) = N(\mathcal{I}) \cdot \sqrt{\Delta_K}$. Using the discriminant, it is possible to bound the value of the shortest vector of an ideal.

Lemma 2.39 ([PR07]). *Let \mathcal{I} be a fractional ideal in a number field K . In any ℓ_p norm for $p \in [1, \infty]$,*

$$n^{1/p} \cdot N(\mathcal{I})^{1/n} \leq \lambda_1(\mathcal{I}) \leq n^{1/p} \cdot N(\mathcal{I})^{1/n} \cdot \Delta_K^{1/(2n)}.$$

Proof. For the upper bound, recall that 2.4 (Minkowski) gives us the bound $\lambda_1(\lambda)_\infty \leq \det(\mathcal{I})^{1/n} = (N(\mathcal{I}) \cdot \sqrt{\Delta_K})^{1/n}$. Using that $\|x\|_p \leq n^{1/p} \cdot \|x\|_\infty$, the bound follows.

For the lower bound, we have that for any element $x \in \mathcal{I}$,

$$\|x\|_p^p = \sum_{i=1}^n |\sigma_i(x)|^p \geq \frac{1}{n} \prod_{i=1}^n (|\sigma_i(x)|^p)^{1/n} = \frac{1}{n} N(x)^{p/n} \geq \frac{1}{n} N(\mathcal{I})^{p/n}.$$

The first inequality uses the arithmetic-geometric mean inequality, and the second uses that $N(x) \geq N(\mathcal{I})$ for any $x \in \mathcal{I}$. The bound follows from taking the p -th root on both sides. \square

Finally, it will be useful to extend the duality that we have for lattices to fractional ideals in \mathcal{O}_K . The importance of dual ideals is such, that the main reductions for Ring-LWE apply for a “dual” version of the decision problem, as we will see in Chapter 4. In this part, we follow [Con].

Definition 2.40. *For any lattice \mathcal{L} in K , the dual of \mathcal{L} is $\mathcal{L}^\vee = \{x \in K : \text{Tr}(x\mathcal{L}) \subset \mathbb{Z}\}$.*

As in the case of general lattices, we have $\mathcal{L}^{\vee\vee} = \mathcal{L}$. For a fractional ideal \mathcal{I} , its dual ideal \mathcal{I}^\vee is also a fractional ideal, as we will see in Lemma 2.41. In \mathcal{O}_K , a particularly interesting dual ideal is \mathcal{O}_K^\vee , called the *codifferent* ideal. First, we have $\text{Tr}(x\mathcal{O}_K) \subset \mathbb{Z}$ for any $x \in \mathcal{O}_K$ (notice that the trace is multiplicative and the trace of any algebraic integer is an integer), hence $\mathcal{O}_K \subset \mathcal{O}_K^\vee$.

Lemma 2.41. *For any fractional ideal \mathcal{I} of \mathcal{O}_K , its dual ideal \mathcal{I}^\vee is also a fractional ideal and satisfies $\mathcal{I}^\vee = \mathcal{I}^{-1}\mathcal{O}_K^\vee$.*

Proof. First, notice that \mathcal{I}^\vee is a finitely generated \mathbb{Z} -module because it is a lattice. According to Definition 2.36, it suffices to check that it is preserved by multiplication by \mathcal{O}_K . Let $a \in \mathcal{I}^\vee$, and observe that $a\mathcal{O}_K \in \mathcal{I}^\vee$ since $\text{Tr}(a\mathcal{O}_K\mathcal{I}) = \text{Tr}(a(\mathcal{O}_K\mathcal{I})) = \text{Tr}(a\mathcal{I}) \subset \mathbb{Z}$.

We will now prove the inclusion to the right. Let $a \in \mathcal{I}^\vee$, and let also $b \in \mathcal{I}$. We have that $\text{Tr}(ab\mathcal{O}_K) \subset \mathbb{Z}$ since $ab \in \mathcal{I}\mathcal{I}^{-1} = \mathcal{O}_K$. Hence, $ab \in \mathcal{O}_K^\vee$. But as $b \in \mathcal{I}$, then $a \in \mathcal{I}^{-1}\mathcal{O}_K^\vee$.

Conversely, take $a \in \mathcal{I}^{-1}\mathcal{O}_K^\vee$, then $a = bc$ where $b \in \mathcal{I}^{-1}$, $c \in \mathcal{O}_K^\vee$. We have that $\text{Tr}(a\mathcal{I}) = \text{Tr}(bc\mathcal{I}) = \text{Tr}(c(b\mathcal{I}))$. As $b \in \mathcal{I}^{-1}$, $b\mathcal{I} \subset \mathcal{O}_K$. But clearly $\text{Tr}(\mathcal{O}_K^\vee\mathcal{O}_K) \subset \mathbb{Z}$, so $\text{Tr}(a\mathcal{I}) = \text{Tr}(c(b\mathcal{I})) \subset \mathbb{Z}$. We conclude that $a \in \mathcal{I}^\vee$. \square

The ideal $(\mathcal{O}_K^\vee)^{-1}$ is called the *different* ideal. This ideal is an integral ideal whose norm is $N((\mathcal{O}_K^\vee)^{-1}) = \Delta_K$. Further details can be found in [Con].

2.2.4 Efficient Computations and CRT

First, recall the well-known Chinese Remainder Theorem in its version for commutative rings (for this work, $R = \mathcal{O}_K$).

Theorem 2.42 (Chinese Remainder Theorem). *Let $\mathcal{I}_1, \dots, \mathcal{I}_r$ be pairwise coprime ideals in a commutative ring R , and let $\mathcal{I} = \mathcal{I}_1 \cdots \mathcal{I}_r$. Then, there is a natural isomorphism $f : R/\mathcal{I} \rightarrow \prod_{i=1}^r (R/\mathcal{I}_i)$ given by $(a \bmod \mathcal{I}) \mapsto (a \bmod \mathcal{I}_1, \dots, a \bmod \mathcal{I}_r)$.*

A subtle aspect is that we are now considering ideals of a ring, and therefore not fractional ideals. Nevertheless, fractional ideals are R -modules, and it is possible to transform them (by multiplication by an algebraic integer) such that they are embedded in R . We must be careful when applying results that hold for ideals in rings, such as the CRT, to situations in which fractional ideals may arise.

A *CRT basis* is a set of elements $c_1, \dots, c_r \in R$ such that $c_i = 1 \bmod \mathcal{I}_i$ and $c_i = 0 \bmod \mathcal{I}_j$ for any $j \neq i$. This basis is used to invert the isomorphism of Theorem 2.42: given $a = (a_1, \dots, a_r) \in \prod_{i=1}^r (R/\mathcal{I}_i)$, then $f^{-1}(a) = \sum_{i=1}^r a_i c_i \bmod \mathcal{I} \in R/\mathcal{I}$.

We now describe how computations over number fields are made practical. Following [LPR10], we say that an algorithm has polynomial complexity (in the context of a number field K) if its running time is polynomial in $n, \log \Delta_K$ and the size of other inputs. Elements $x \in K$ can be represented in a basis $\{b_1, \dots, b_n\}$ for \mathcal{O}_K : if the coefficients are integers, $x \in \mathcal{O}_K$, otherwise they will be rationals. Hence, n rational coordinates suffice to represent

any vector. Moreover, integral ideals \mathcal{I} are also represented by a \mathbb{Z} -basis of algebraic integers, so we can use the same basis as before. A fractional ideal \mathcal{I} can be represented as an ideal plus a denominator $d \in \mathcal{O}_K$ such that $d\mathcal{I} \subset \mathcal{O}_K$.

Proposition 2.43 (Efficient computations [LPR10], [LPR13]). *Let K be a number field, let \mathcal{O}_K be its ring of integers, and let \mathcal{I}, \mathcal{J} be fractional \mathcal{O}_K -ideals. There exist algorithms that perform the following computations efficiently:*

- *Field: addition, subtraction, product and inversion of elements in K and \mathcal{O}_K .*
- *Ideal: check that a basis generates \mathcal{I} , compute $N(\mathcal{I})$, compute \mathcal{I}^{-1} and \mathcal{I}^\vee , reduce $a \bmod \mathcal{I}$ for $a \in K$. For two ideals: compute $\mathcal{I} \cdot \mathcal{J}$, sample from \mathcal{I}/\mathcal{J} if $\mathcal{J} \subset \mathcal{I}$.*
- *Sampling: obtain samples from $D_{\mathbf{r}}$ over $K_{\mathbb{R}}$.*
- *CRT: Obtain a CRT basis for integral ideals $\mathcal{I}_1, \dots, \mathcal{I}_r \subset \mathcal{O}_K$.*

Let \mathcal{I}, \mathcal{J} be fractional \mathcal{O}_K -ideals. We now prove two results that yield an efficiently computable isomorphism (of \mathcal{O}_K -modules) between $\mathcal{I}/q\mathcal{I}$ and $\mathcal{J}/q\mathcal{J}$. These will be needed for the reductions in Chapter 4. Again, we follow [LPR10].

Lemma 2.44. *For any two integral ideals $\mathcal{I}, \mathcal{J} \subset \mathcal{O}_K$, there exists a $t \in \mathcal{I}$ such that $t\mathcal{I}^{-1} \subset \mathcal{R}$ is an ideal coprime to \mathcal{J} . Such t can be found efficiently given \mathcal{I} and the prime factorization of \mathcal{J} .*

Proof. Suppose that $\mathcal{J} = \mathfrak{p}_1^{d_1} \cdots \mathfrak{p}_r^{d_r}$, where $d_i > 0$, is the prime factorization of \mathcal{J} . Then, we can write $\mathcal{I} = \mathcal{I}' \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r}$, where \mathcal{I}' is coprime with \mathcal{J} and $0 \leq e_i \leq d_i$. The exponents e_i can be found efficiently by binary searching on the greatest exponent of \mathfrak{p}_i that divides \mathcal{I} ; notice that e_i cannot be greater than $\log(N(\mathcal{I}))/\log(N(\mathfrak{p}_i))$. Now, we can take $t_i \in \mathfrak{p}_i^{e_i}$ and $t_i \notin \mathfrak{p}_i^{e_i+1}$ for each i . Since \mathcal{I}' and \mathfrak{p}_i are coprime for each i , we can use the CRT to find a $t \in \mathcal{O}_K$ such that $t \equiv 0 \bmod \mathcal{I}'$ and $t \equiv t_i \bmod \mathfrak{p}_i^{e_i+1}$ (this can also be done efficiently using a CRT basis). Besides, $t \in \mathcal{I}$ as $t \in \mathcal{I}'$ and $t \equiv 0 \bmod \mathfrak{p}_i^{e_i}$ for every i .

Notice that $t\mathcal{I}^{-1}$ is an integral ideal, since for $t \in \mathcal{I}$, $a \in \mathcal{I}^{-1}$, elements of the form $ta \in \mathcal{I}\mathcal{I}^{-1} = \mathcal{O}_K$. It remains to prove that $t\mathcal{I}^{-1}$ is coprime with \mathcal{J} , or equivalently, that no \mathfrak{p}_i divides $t\mathcal{I}^{-1}$. Suppose that \mathfrak{p}_i divides $t\mathcal{I}^{-1}$. Then, $\mathfrak{p}_i\mathcal{I} \mid (t)$, and since $\mathfrak{p}_i^{e_i+1} \mid \mathfrak{p}_i\mathcal{I}$ we have that $t \in \mathfrak{p}_i^{e_i+1}$, which contradicts that $t \not\equiv 0 \bmod \mathfrak{p}_i^{e_i+1}$. \square

Lemma 2.45 (Clearing ideals). *Let $\mathcal{I}, \mathcal{J} \subset \mathcal{O}_K$ and $t \in \mathcal{I}$ as in Lemma 2.44. Let \mathcal{M} be a fractional \mathcal{O}_K -ideal. The function $\theta_t : \mathcal{M} \rightarrow \mathcal{I}\mathcal{M}$ given by $x \mapsto tx$ induces an isomorphism of \mathcal{O}_K -modules from $\mathcal{M}/\mathcal{J}\mathcal{M}$ to $\mathcal{I}\mathcal{M}/\mathcal{I}\mathcal{J}\mathcal{M}$. This isomorphism can be efficiently inverted given $\mathcal{I}, \mathcal{J}, \mathcal{M}$, and t .*

Proof. Consider the map $\tilde{\theta}_t : \mathcal{M} \rightarrow \mathcal{I}\mathcal{M}/\mathcal{I}\mathcal{J}\mathcal{M}$ given by $\tilde{\theta}_t(x) = tx \bmod \mathcal{I}\mathcal{J}\mathcal{M}$. It clearly is an homomorphism of \mathcal{O}_K -modules since it is a multiplication by an element of the ring $t \in \mathcal{O}_K$. First, we prove that $\ker(\tilde{\theta}_t) = \mathcal{J}\mathcal{M}$. Let $x \in \ker(\tilde{\theta}_t)$, then $tx \in \mathcal{I}\mathcal{J}\mathcal{M}$. Hence, $(t\mathcal{I}^{-1})(x\mathcal{M}^{-1}) \subset \mathcal{J}$, and since $t\mathcal{I}^{-1}$ and \mathcal{J} are coprime, necessarily $x\mathcal{M}^{-1} \subset \mathcal{J}$ and $x \in \mathcal{J}\mathcal{M}$. On the other direction, if $x \in \mathcal{J}\mathcal{M}$, then $\tilde{\theta}_t(x) = tx \in \mathcal{I}\mathcal{J}\mathcal{M}$.

If we prove that $\tilde{\theta}_t$ is surjective, we will be done by the first isomorphism theorem. Let $y \in \mathcal{I}\mathcal{M}$, we want to find $x \in \mathcal{M}$ such that $\tilde{\theta}_t(x) \equiv y \bmod \mathcal{I}\mathcal{J}\mathcal{M}$. Since $t\mathcal{I}^{-1}$ and \mathcal{J} are coprime, we can (efficiently) find an element $c \in t\mathcal{I}^{-1}$ such that $c \equiv 1 \bmod \mathcal{J}$. Now, take $a = cy \in t\mathcal{I}^{-1}\mathcal{I}\mathcal{M} = t\mathcal{M}$ and let $x = a/t \in \mathcal{M}$. Clearly, we have that $\tilde{\theta}_t(x) = a$. Besides, $a - y = y(c - 1) \in \mathcal{I}\mathcal{J}\mathcal{M}$, so $a \equiv y \bmod \mathcal{I}\mathcal{J}\mathcal{M}$ and we conclude that $\tilde{\theta}_t(x) \equiv y \bmod \mathcal{I}\mathcal{J}\mathcal{M}$ as desired. The proof is constructive, showing that the isomorphism can be inverted efficiently. \square

For some results, such as in [RSW18], it is useful that the norm of $t \in \mathcal{I}$ in Theorem 2.45 is small, in order to avoid large noise amplification. It is worth mentioning that t can be chosen in a different way, such as using Gaussian samples, but this will not be addressed in this work.

2.2.5 Ideal Lattice Problems

All computational problems for lattices in Section 2.1.3 can be formulated for ideal lattices, where \mathcal{L} is replaced by an ideal \mathcal{I} (notice that considering only integral ideals is sufficient as we can transform any fractional ideal into an integral ideal). When this is the case, we append the prefix R - to the problem: R -SVP, R -BDD $_{\mathcal{I},d}$, etc.

A computational problem may be easy for ideal lattices but hard for general lattices. An example is the R -GapSVP $_{\gamma}$ problem, which is easy for ideal lattices (for certain parameters) but believed to be hard in a general setting. The reason behind this is that ideal lattices are a special case of general lattices, so a hardness result that relies on an assumption on ideal lattices is weaker than a result on general lattices. This is exactly what happens with LWE and Ring-LWE; as we will see in Chapters 3 and 4, the hardness results for Ring-LWE are similar to the results for LWE, but on ideal lattices only.

Chapter 3

Learning with Errors

In this chapter, we introduce Learning with Errors formally, following the approach taken in the seminal work by Regev [Reg09]. We present the ideas behind the worst-case reduction from standard lattice problems (GapSVP, SIVP) to LWE, including many of the proofs but focusing on intuition. At the end of the chapter, we introduce the techniques in [PRSD17] for finding the center of an *oracle*, an important step in proving the hardness of Ring-LWE. Nevertheless, we will not require any concepts in algebraic number theory and ideal lattices.

3.1 Overview

The Learning with Errors problem can be found in two main versions. Both of them are equivalent, as proven in [Reg09]. We will mostly consider the continuous version due to its similarities with Ring-LWE, however, some results will be proven for the discrete case for simplicity and similarity with applications. The parameters of a LWE instance are two positive integers n, q and an error distribution ϕ over $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ (or χ over \mathbb{Z}_q , for the discrete case) where n is the vector size and q is the field size. The formal definition of the LWE distribution is the following [Reg09].

Definition 3.1 (LWE distribution). *For a secret $\mathbf{s} \in \mathbb{Z}_q^n$, the continuous LWE distribution $A_{\mathbf{s}, \phi}$ on $\mathbb{Z}_q^n \times \mathbb{T}$ is the probability distribution obtained by choosing a vector $\mathbf{a} \in \mathbb{Z}_q^n$ uniformly at random, sampling $e \in \mathbb{T}$ from ϕ , and outputting $(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle / q + e \bmod 1)$.*

The discrete LWE distribution $A_{\mathbf{s}, \chi}$ on $\mathbb{Z}_q^n \times \mathbb{Z}_q$ is the probability distribution obtained by choosing a vector $\mathbf{a} \in \mathbb{Z}_q^n$ uniformly at random, sampling $e \in \mathbb{Z}_q$ from χ , and outputting $(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e \bmod q)$.

A probabilistic algorithm which samples from $A_{\mathbf{s}, \chi}$ will be called an *LWE-oracle*. Most commonly, the error distributions will be Gaussian, as defined in Section 2.1.2. For the continuous case and for a width parameter β , this is simply $\phi_\beta = D_\beta \bmod 1$. In the case of a discrete distribution over \mathbb{Z}_q^n , it is the discrete Gaussian $\chi_\beta = D_{\mathbb{Z}, q, \beta}$. We will use this notation from now on.

Similarly to lattice problems, LWE has decision and search variants (the definitions for the discrete versions are analogous). In both cases, the secret \mathbf{s} is fixed for all samples, and n (dimension) and m (number of samples) parametrize the problem.

Definition 3.2 (Search-LWE, SLWE $_{\phi, q}$). *Let $\mathbf{s} \in \mathbb{Z}_q^n$ be a uniformly random vector. Given m independent samples $(\mathbf{a}_i, b_i) \in \mathbb{Z}_q^n \times \mathbb{T}$ drawn from $A_{\mathbf{s}, \phi}$, find \mathbf{s} .*

Definition 3.3 (Average-case Decision-LWE, LWE $_{\phi, q}$). *Let $\mathbf{s} \in \mathbb{Z}_q^n$ be a uniformly random vector. Given m independent samples $(\mathbf{a}_i, b_i) \in \mathbb{Z}_q^n \times \mathbb{T}$ either (1) drawn from $A_{\mathbf{s}, \phi}$ or (2) drawn from the uniform distribution, determine which is the case with non-negligible advantage.*

The definitions presented above might hide the actual structure of the LWE problem, which is solving a system of linear equations with noise. Consider the discrete version, and suppose that we are given $m \geq n$ samples of LWE $(\mathbf{a}_1, b_1), \dots, (\mathbf{a}_m, b_m)$ where $\mathbf{a}_i \in \mathbb{Z}_q^n$ and the error distribution is 0. Then, the $b_i = \langle \mathbf{a}_i, \mathbf{s} \rangle$ and the search problem reduces to solving $A\mathbf{s} = \mathbf{b}$, where A is an $m \times n$ matrix whose rows are the \mathbf{a}_i . If the rank of A is n (which happens with high probability as the \mathbf{a}_i are sampled from uniform), then it is straightforward to find \mathbf{s} using Gaussian elimination. LWE only becomes harder when we add a small noise e_i to the b_i . A natural interpretation of decision-LWE is that the b_i are pseudo-random values generated from the LWE distribution.

A nice property of LWE is that the search and decision problems are in fact equivalent, i.e., if one has access to an efficient distinguisher for Decision-LWE, a polynomial-time algorithm for Search-LWE can be constructed. Unfortunately, the result cannot be extended to Ring-LWE¹.

Lemma 3.4 (Search to Decision reduction [Reg09]). *Let $n \geq 1$ be some integer, $2 \leq q \leq \text{poly}(n)$ a prime, $\epsilon = \epsilon(n)$ a negligible amount and χ some distribution on \mathbb{Z}_q . Given an algorithm W that accepts with probability greater than $1 - \epsilon$ on inputs from $A_{\mathbf{s}, \chi}$ and rejects on inputs from the uniform distribution also with probability larger than $1 - \epsilon$, there exists an efficient algorithm that, given samples of $A_{\mathbf{s}, \chi}$, outputs \mathbf{s} with probability $1 - \text{poly}(\epsilon)$.*

Proof. Consider samples (\mathbf{a}, b) from $A_{\mathbf{s}, \chi}$. We will find the first coordinate s_1 of \mathbf{s} using the output of W . The remaining coordinates can be found similarly. Let $k \in \mathbb{Z}_q$ and consider the transformation $f_l : \mathbb{Z}_q^n \times \mathbb{Z}_q \rightarrow \mathbb{Z}_q^n \times \mathbb{Z}_q$ given by $f_l(\mathbf{a}, b) = (\mathbf{a} + (l, 0, \dots, 0), b + lk)$ where l is chosen uniformly at random on \mathbb{Z}_q . Notice that f_l takes the uniform distribution into itself, since it is a (coordinate-wise) linear map over a finite field. If we transform a sample from $A_{\mathbf{s}, \chi}$, we obtain (modulo q):

$$f_l(\mathbf{a}, b) = f_l(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e) = (\mathbf{a} + (l, 0, \dots, 0), \langle \mathbf{a}, \mathbf{s} \rangle + e + kl).$$

The distribution of the output depends on the value of k . We have two possibilities:

- If $k = s_1$, we have that $\langle \mathbf{a} + (l, 0, \dots, 0), \mathbf{s} \rangle = \langle \mathbf{a}, \mathbf{s} \rangle + s_1 l$. Hence, f_l preserves the distribution $A_{\mathbf{s}, \chi}$.
- If $k = s_1 + c$ where $c \neq 0$, then $\langle \mathbf{a} + (l, 0, \dots, 0), \mathbf{s} \rangle = \langle \mathbf{a}, \mathbf{s} \rangle + s_1 l + cl$. As l is uniformly random and q is prime, $cl \bmod q$ is uniformly distributed. Thus the output of f_l is also uniformly distributed.

Hence, we can use the distinguisher W to test whether $k = s_1$, coordinate-wise. Since there are q possibilities and n coordinates, we try all the possibilities in $\text{poly}(n)$. \square

In the previous result, it suffices to have a non-negligible advantage in the distinguisher; we just need to iterate on several values for l to increase its advantage [Ste14]. Another property of LWE is that worst-case hardness implies average-case hardness, meaning that if the problem is hard for a non-negligible fraction of secrets \mathbf{s} (i.e., the worst-case instances), then it is also hard for random choices of \mathbf{s} (average-case). This is a necessary property in most cryptographic constructions - otherwise, keys could not be chosen at random (among other implications).

¹A search-to-decision reduction for Ring-LWE appears in [Ros20], using techniques from [PRSD17] that we introduce in Chapter 4

Lemma 3.5 (Average-case to worst-case reduction [Reg10]). *Let $n, q \geq 1$ be integers and χ a distribution on \mathbb{Z}_q . Given an efficient algorithm W that distinguishes $A_{\mathbf{s}, \chi}$ from uniform for a non-negligible fraction of $\mathbf{s} \in \mathbb{Z}_q^n$, there exists an algorithm W' that for any \mathbf{s} accepts on inputs from $A_{\mathbf{s}, \chi}$ with probability exponentially close to 1 and rejects on inputs from the uniform distribution with probability exponentially close to 1.*

Proof Sketch. We give the idea of the proof; the calculation of the specific probabilities and number of iterations can be found in [Reg09]. Let $\mathbf{t} \in \mathbb{Z}_q^n$ and consider the map $f_{\mathbf{t}} : \mathbb{Z}_q^n \times \mathbb{Z}_q \rightarrow \mathbb{Z}_q^n \times \mathbb{Z}_q$ given by

$$f_{\mathbf{t}}(\mathbf{a}, b) = (\mathbf{a}, b + \langle \mathbf{a}, \mathbf{t} \rangle).$$

Notice that this function transforms $A_{\mathbf{s}, \chi}$ into $A_{\mathbf{s} + \mathbf{t}, \chi}$ and the uniform distribution into itself. The algorithm W' proceeds as follows: on inputs of the form (\mathbf{a}, b) , sample \mathbf{t} uniformly at random and estimate the probability acceptance of W on both (\mathbf{a}, b) and $f_{\mathbf{t}}(\mathbf{a}, b)$. This is made by running W a polynomial number of times. If the probabilities differ noticeably, W' accepts. Otherwise, a different \mathbf{t} is sampled and the process is repeated. If the probabilities did not differ after a sufficient (polynomial) number of iterations, W' rejects. \square

We present the main hardness result for LWE below. The theorem, due to Regev [Reg09], is a quantum reduction from worst-case GapSVP_{γ} and SIVP_{γ} to decision-LWE on general lattices. A quantum reduction means that an efficient distinguisher for LWE samples would yield an efficient quantum algorithm for solving *any instance* of GapSVP_{γ} , SIVP_{γ} . The reduction transforms lattice problem instances into LWE instances, that can be solved using the distinguisher. This reduction is meaningful since these problems are believed to be hard for quantum computers. As of 2021, no polynomial-time algorithm that achieves polynomial approximation factors for GapSVP_{γ} , SIVP_{γ} on general lattices is known.

Theorem 3.6 (Main theorem, extended version). *For any number of samples $m = \text{poly}(n)$, any prime $q \leq 2^{\text{poly}(n)}$, and any discrete Gaussian distribution χ_{α} such that $\alpha q > 2\sqrt{n}$ and $0 < \alpha < 1$, if there is an efficient algorithm that solves $\text{LWE}_{\chi, q}$, then there exists an efficient quantum algorithm that solves GapSVP_{γ} , SIVP_{γ} for $\gamma = \tilde{O}(n/\alpha)$ on n -dimensional lattices.*

In applications of LWE, the choice of the modulus q will depend on the dimension n and generally be polynomial. For example, the cryptosystem in [Reg09] chooses $n^2 \leq q \leq 2n^2$, and Frodo KEM [BCD⁺16] uses $n = 752 \approx 2^{9.5}$ and $q = 2^{15}$ (the requirement of q being prime is not necessary, as we will see later). If we choose to have a small q , then α needs to be larger; this makes γ smaller and hence the lattice problems become harder (since the approximation factor decreases). Intuitively, a larger α results in a wider error distribution and therefore in a harder LWE.

Making the reduction entirely classical remains as one of the most important open problems surrounding LWE. There has been partial progress; in [BLP⁺13], a reduction that achieves subexponential approximation factors is presented, but at the cost of squaring the dimension of the LWE problem to be solved.

3.2 Regev's Approach

Our goal in this section is to explain the main components of the reduction to LWE, with a special emphasis on the parts that apply to Ring-LWE. The first thing to note is that Theorem 3.6 is in fact a generalization of the following theorem, which is the main result in [Reg09].

Theorem 3.7 (Main theorem). *Let $q \geq 2$, $\alpha \in (0, 1)$ such that $\alpha q > 2\sqrt{n}$. There is a polynomial time quantum reduction from $\text{DGS}_{\sqrt{2n}\eta_\epsilon(\mathcal{L})/\alpha}$ (for a negligible $\epsilon = \epsilon(n) > 0$) to $\text{SLWE}_{q, \phi_\alpha}$.*

Theorems 3.6 and 3.7 are very similar. The main differences are the reduction to search-LWE instead of decision-LWE, the use of the continuous LWE distribution, and the lattice problem where we reduce from. To obtain Theorem 3.6 from Theorem 3.7, we therefore combine the following problem equivalences.

1. *Search \iff Decision.* The reduction in Lemma 3.4 implies that the result on SLWE extends directly to LWE (for a prime q). We remark that one of the main difficulties in the ring setting is the lack of such a self-reducibility result for Ring-LWE.
2. *Discrete \iff Continuous.* Handling discrete error samples does not make the problem easier or harder ([Reg09]).
3. *Lattice problems.* There exist tight reductions from GapSVP_γ and SIVP_γ to DGS_φ ; these were mentioned in Section 2.1.3. Notice that DGS_φ is a natural problem to work with when handling discrete Gaussian distributions.

Hence, we focus on proving Theorem 3.7. The reduction is based on repeated applications of a procedure that takes some samples of $D_{\mathcal{L}, r}$ and outputs the same number of samples of $D_{\mathcal{L}, r'}$ for a smaller $r' \leq r/2$. We refer to this procedure as the Iterative Step (IS), which is also used (with minor changes) in the reduction to Ring-LWE. In order to reduce the norm of the samples, the IS uses a quantum algorithm that makes calls to an oracle for SLWE. The connection with Theorem 3.7 is that, if we start sampling from a wide Gaussian distribution (recall the bootstrapping procedure in Lemma 2.14), and we repeat the IS a sufficient number of times, the samples will eventually come from a narrow distribution, which is equivalent to solving DGS_φ . Let us formalize this argument.

Proposition 3.8 (Iterative Step). *Let $q \geq 2$, $\alpha \in (0, 1)$, and \mathcal{O} be an oracle that solves $\text{SLWE}_{q, \phi_\alpha}$ on input a polynomial number of samples. There exists an efficient quantum algorithm that, given access to \mathcal{O} , an n -dimensional lattice \mathcal{L} , a parameter $r > \sqrt{2}q\eta_\epsilon(\mathcal{L})$ for some negligible $\epsilon = \epsilon(n) > 0$, and a polynomial number of samples from $D_{\mathcal{L}, r}$, outputs the same number of samples from $D_{\mathcal{L}, r\sqrt{n}/(\alpha q)}$.*

Proof of Theorem 3.7

Let \mathcal{L} be an n -dimensional lattice and let $r > \sqrt{2n}\eta_\epsilon(\mathcal{L})/\alpha$. We want to obtain a sample from $D_{\mathcal{L}, r}$ by repeated applications of the IS (Proposition 3.8). Define the sequence $r_{i+1} = r_i \cdot (\alpha q/\sqrt{n})$ starting from $r_0 = r$, hence $r_i = (\alpha q/\sqrt{n})^i$. Notice that $(\alpha q/\sqrt{n}) > 2$ by assumption, thus the sequence is increasing.

First, notice that we can obtain samples from a wide distribution efficiently. By Lemmas 2.6 and 2.12, it follows that² $r_{3n} > r2^{3n} > 2^{2n}\lambda_n(\mathcal{L})$, so we can apply the bootstrapping method from Lemma 2.14 to sample from $D_{\mathcal{L}, r_{3n}}$.

To reduce the norm of our samples, notice that if we have n^c samples from $D_{\mathcal{L}, r_{i+1}}$, we can obtain n^c samples from $D_{\mathcal{L}, r_i}$ by applying the IS. The condition on r is always satisfied as, for any $i \geq 1$, we have that $r_i \geq r_1 = r\alpha q/\sqrt{n} > \sqrt{2}\eta_\epsilon(\mathcal{L})\alpha q$ by assumption. Finally, notice that applying the IS on r_1 yields samples from $D_{\mathcal{L}, r}$. We conclude that it suffices to bootstrap n^c samples from $D_{\mathcal{L}, r_{3n}}$ and to apply the IS for $i = 3n, \dots, 1$. \square

²This is a gross bound. One could find a better estimation, but we do not focus on making the reduction implementable in practice.

3.2.1 The Iterative Step

The only thing that remains is the construction of the Iterative Step, which is divided into two parts as presented in Figure 3.1. The first part is a classical procedure that solves a BDD problem on the dual lattice \mathcal{L}^* using a LWE oracle and samples from $D_{\mathcal{L},r}$. The second part is a quantum reduction that uses a BDD oracle on \mathcal{L}^* to create samples from a discrete Gaussian distribution on \mathcal{L} . If both parts are applied sequentially, the width of the distribution of the samples is reduced. The samples that are generated are *independent*, meaning that there is no correlation between them (in other words, the probability that m samples obtained in this way are linearly dependent is the same as for m uniformly random samples).

We present both results separately. For the quantum part, we only introduce the main result, which will be used in Chapter 4 without change. On the other hand, we present the classical part in detail. The ideas behind the classical part for LWE and for Ring-LWE are very similar, but they differ in the technical component. Broadly speaking, this is because when dealing with ring elements, we have less control on the behaviour of the lattice vectors under the canonical embedding.

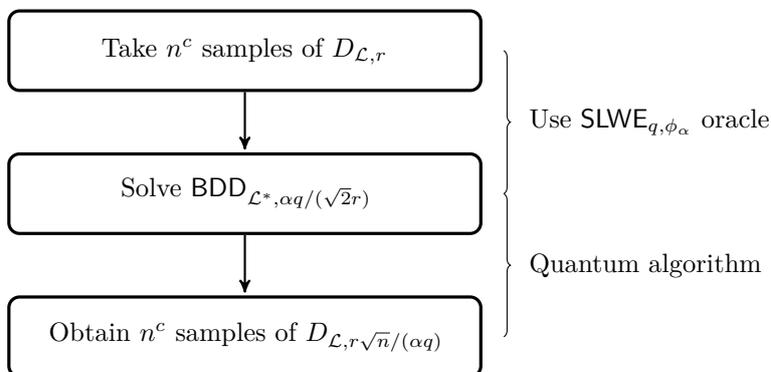


Figure 3.1: Diagram of the Iterative Step

Proposition 3.9 (Classical part of the IS). *Let $\epsilon = \epsilon(n) > 0$ be negligible, $q \geq 2$, $\alpha \in (0, 1)$. Given access to an oracle that solves $\text{SLWE}_{q, \phi_\alpha}$ given a polynomial number of samples, there exists an efficient algorithm that, for any n -dimensional lattice \mathcal{L} , $r > \sqrt{2}q\eta_\epsilon(\mathcal{L})$, solves $\text{BDD}_{\mathcal{L}^*, \alpha q / (\sqrt{2}r)}$ using a polynomial number of samples from $D_{\mathcal{L},r}$.*

Notice that the parameters that appear in Proposition 3.9 determine those appearing in Proposition 3.8. In particular, the condition $r > \sqrt{2}q\eta_\epsilon(\mathcal{L})$ is imposed by the classical step and is related to the need to sum discrete and continuous Gaussians. For everything to behave nicely, we need to work above the smoothing parameter (see Lemma 2.9) and below.

Proposition 3.10 (Quantum part of the IS). *Given any n -dimensional lattice \mathcal{L} , a real $g < \lambda_1(\mathcal{L}^*)/2$, a vector \mathbf{r} such that $r_i \geq 1$ for $i = 1, \dots, n$, and an oracle that solves $\text{GDP}_{\mathcal{L}^*, g}$ or BDD_g with overwhelming probability, there is an efficient quantum algorithm that outputs an independent sample from $D_{\mathcal{L}, \mathbf{r}\sqrt{n}/(\sqrt{2}g)}$.*

Alternatively, if $g < \lambda_1(\mathcal{L}^)/(\sqrt{8n})$, then the algorithm outputs an independent sample from $D_{\mathcal{L}, \mathbf{r}/(2g)}$.*

The version of Proposition 3.10 presented here is a generalization, introduced in [PRSD17], of the original reduction in [Reg09]. There are two differences between them. The first is that in [Reg09], the vector \mathbf{r} does not appear, but rather the samples are from a discrete spherical Gaussian distribution $D_{\mathcal{L},r}$ where $r = 1$. The extension to elliptical Gaussians with

vectors \mathbf{r} works since each $r_i \geq 1$. Hence, it is possible to apply a suitable transformation that yields samples from a (wider) discrete elliptical Gaussian on each coordinate. The second difference is that in [Reg09], the oracle solves BDD, and GDP is not mentioned. However, this fact is intrinsic to the original reduction, as the BDD oracle is called on points whose errors are drawn from a Gaussian distribution (which is precisely the definition of GDP). The equivalence with BDD follows by an argument similar to Lemma 2.13, which is that the probability of finding long vectors when sampling from a Gaussian is very small. The oracle is allowed to be mistaken with negligible probability, so this is not an issue.

The proof of the Iterative Step follows directly from both parts. Starting with n^c samples of $D_{\mathcal{L},r}$, we can use the $\text{SLWE}_{q,\phi_\alpha}$ oracle to construct an efficient algorithm that solves $\text{BDD}_{\mathcal{L}^*,g}$ where $g = \alpha q / (\sqrt{2}r)$ by Proposition 3.9. Then, we can use such algorithm as the oracle in Proposition 3.10. Running it n^c times and setting $\mathbf{r} = \mathbf{1}$ (a spherical Gaussian), we obtain n^c independent random samples of $D_{\mathcal{L},r'}$, where $r' = \sqrt{n}/(\sqrt{2}g)$. Plugging in the value of g yields $r' = r\sqrt{n}/(\alpha q)$ as desired.

3.2.2 The Classical Part

We will now present the insights of the classical part of the Iterative Step (Proposition 3.9). Our goal is to, starting from a lattice \mathcal{L} , a number of samples from $D_{\mathcal{L},r}$, and an oracle that solves SLWE, construct an algorithm that solves $\text{BDD}_{\mathcal{L},d}$ in the dual lattice for a suitable d . The main idea (Lemma 3.12) is to formulate BDD as an SLWE instance and then solve it by calling the oracle. Recall that $\text{BDD}_{\mathcal{L}^*,d}$ consists in finding a lattice point $\mathbf{x} \in \mathcal{L}^*$ given a perturbed vector $\mathbf{y} = \mathbf{x} + \mathbf{e}$, for some bounded error $\|\mathbf{e}\| \leq d$. We will create SLWE instances where the secret \mathbf{s} is related to the target lattice point \mathbf{x} ; once \mathbf{s} is found with the oracle, we will be able to recover \mathbf{x} .

To carry out the proof, we will benefit from two additional facts. First, it will suffice to have a SLWE oracle for a single error distribution ϕ_α . This makes the result stronger and also allows for a certain slack in the error distribution of the samples, since a smaller error than expected is not problematic (Lemma 3.11). Second, it will suffice to solve a modular (and easier) version of BDD (Lemma 3.13). We present the three Lemmas and prove Proposition 3.9 below.

Lemma 3.11. *Given access to an oracle that solves $\text{SLWE}_{q,\phi_\alpha}$, we can efficiently solve $\text{SLWE}_{q,\phi_\beta}$ for any (unknown) $\beta \leq \alpha$ with overwhelming probability.*

Proof Sketch. The idea is to transform the samples from A_{s,ϕ_β} into samples from A_{s,ϕ_α} . This is done by adding extra noise to the samples. Notice that if an error term e (not related to the \mathbf{e} from the BDD instance) is sampled from ϕ_β , then $e_0 = e + e'$ will correspond to a sample of ϕ_α whenever e' is sampled from $\phi_{\sqrt{\alpha^2 - \beta^2}}$.

More specifically, given an oracle W for $\text{SLWE}_{q,\phi_\alpha}$, the procedure consists of adding error terms e' sampled from different ϕ_γ and testing W on such inputs. If $\gamma^2 \approx \alpha^2 - \beta^2$, then the statistical distance between ϕ_α and $\phi_\beta + \phi_\gamma$ will be very small, hence the probability that the output of W is correct will be large. To find such a γ , it suffices to test a sufficient (polynomial) number of possible $\gamma \in [0, \alpha]$. \square

Lemma 3.12. *Let $\text{BDD}_{\mathcal{L}^*,d}$ be instance given by a vector $\mathbf{y} = \mathbf{x} + \mathbf{e}$ such that $\mathbf{x} \in \mathcal{L}^*$ and $\|\mathbf{e}\| \leq d$. There is a probabilistic algorithm that, given such instance, an $r \geq \sqrt{2}q\eta(\mathcal{L})$, and a sample from $D_{\mathcal{L},r}$, outputs a sample from A_{s,ϕ_β} where $\beta = \sqrt{(r\|\mathbf{e}\|/q)^2 + (rd/q)^2}$ and $\mathbf{s} = (\mathcal{L}^*)^{-1}\mathbf{x} \bmod q$.*

Proof Sketch. Let $\mathbf{v} \in \mathcal{L}$ be a sample from $D_{\mathcal{L},r}$, and $\mathbf{a} \in \mathbb{Z}_q^n$ its coefficient vector such that $\mathcal{L}\mathbf{a} = \mathbf{v} \bmod q$ (notice that we are abusing notation, writing \mathcal{L} for a matrix basis of the lattice,

in order to emphasize duality). Let e' be a sample from $\phi_{\alpha/(2\sqrt{\pi})}$ for some $\alpha \geq \beta$. We claim that

$$(\mathbf{a}, \langle \mathbf{y}, \mathbf{v} \rangle / q + e' \bmod 1)$$

is a sample from A_{s, ϕ_β} . Taking different \mathbf{v} and e' will lead to a set of independent samples.

To see why this is true, we first notice that the distribution of \mathbf{a} is statistically close to uniform (differing by a negligible amount). The coefficient vector \mathbf{a} is calculated mod q , hence it suffices if it has a uniform distribution mod q . Informally, we are sampling \mathbf{v} above the smoothing parameter of \mathcal{L} , so Lemma 2.9 tells us that our distribution is uniform in $\mathbb{R}^n / \mathcal{L}$. To extend this mod q , we require our radius to be a factor q larger than η_ϵ (and therefore the condition $r \geq \sqrt{2}q\eta_\epsilon(\mathcal{L})$).

To study the second element, we can condition on a fixed value of \mathbf{a} , and observe that

$$\langle \mathbf{y}, \mathbf{v} \rangle / q + e' \bmod 1 = \langle \mathbf{x}, \mathbf{v} \rangle / q + \langle \mathbf{e}/q, \mathbf{v} \rangle + e' \bmod 1$$

For the first term,

$$\langle \mathbf{x}, \mathbf{v} \rangle = (\mathcal{L}^*)^{-1} \langle \mathbf{x}, \mathbf{v} \rangle \mathcal{L}^* = \langle (\mathcal{L}^*)^{-1} \mathbf{x}, \mathcal{L}^{-1} \mathbf{v} \rangle = \langle \mathbf{s}, \mathbf{a} \rangle \bmod q,$$

since the $\mathcal{L}^{-1} = (\mathcal{L}^*)^T$ and the coefficient vectors \mathbf{s}, \mathbf{a} are integer vectors. Finally, one can prove using Lemma 2.10 that the distribution of $\langle \mathbf{e}/q, \mathbf{v} \rangle + e'$ is statistically close to ϕ_β , as claimed. \square

For the third lemma, let $q\text{-BDD}_{\mathcal{L}, d}$ be the problem in which, given a $\text{BDD}_{\mathcal{L}, d}$ instance with solution $x \in \mathcal{L}$, the goal is to find $x \bmod q\mathcal{L}$. We prove that solving $q\text{-BDD}_{\mathcal{L}, d}$ is sufficient to solve $\text{BDD}_{\mathcal{L}, d}$ on the same lattice.

Lemma 3.13. *Given access to an oracle that solves $q\text{-BDD}_{\mathcal{L}, d}$ for a lattice \mathcal{L} and an integer $q \geq 2$, we can solve $\text{BDD}_{\mathcal{L}, d}$ efficiently on the same lattice.*

Proof. Let B be a basis of \mathcal{L} , and let \mathcal{O} be a $q\text{-BDD}_{\mathcal{L}, d}$ oracle. On input $(\mathcal{L}, d, \mathbf{y})$ where $d < \lambda_1(\mathcal{L})/2$, $\mathbf{y} = \mathbf{x} + \mathbf{e}$ for $\mathbf{x} \in \mathcal{L}$, $\|\mathbf{e}\| \leq d$, the oracle \mathcal{O} outputs a coefficient vector \mathbf{a} such that $B\mathbf{a} = \mathbf{x} \bmod q$.

We start with a point $\mathbf{y} = \mathbf{x} + \mathbf{e}$ close to \mathcal{L} , and define a sequence $\mathbf{y}_1, \mathbf{y}_2, \dots$ (where $\mathbf{y}_1 = \mathbf{y}$) as follows. Let \mathbf{a}_i be the coefficient vector of the lattice point $\mathbf{x}_i \in \mathcal{L}$ closest to \mathbf{y}_i , and define $\mathbf{y}_{i+1} = (\mathbf{y}_i - B(\mathbf{a}_i \bmod q))/q$. We claim that the nearest lattice point to \mathbf{y}_{i+1} is $\mathbf{x}_{i+1} = B(\mathbf{a}_i - (\mathbf{a}_i \bmod q))/q \in \mathcal{L}$. It is indeed a lattice point since q divides each coefficient of $\mathbf{a}_i - (\mathbf{a}_i \bmod q)$. To see that it is the nearest point, notice that

$$|\mathbf{y}_{i+1} - \mathbf{x}_{i+1}| = \left| \frac{\mathbf{y}_i - B(\mathbf{a}_i \bmod q) - B(\mathbf{a}_i - (\mathbf{a}_i \bmod q))}{q} \right| = \left| \frac{\mathbf{y}_i - \mathbf{x}_i}{q} \right| < \frac{d}{q^i}.$$

Each of the \mathbf{y}_i can be computed using \mathcal{O} , since it suffices to find $\mathbf{a}_i \bmod q$. Hence, after n steps, the point \mathbf{y}_{n+1} is at distance at most $d/q^n < \lambda_1(\mathcal{L})/(2q^n)$ from the lattice. Using Lemma 2.23, we can recover \mathbf{x}_{n+1} as the nearest point in the lattice to \mathbf{y}_{n+1} , and hence $\mathbf{a}_{n+1} = B^{-1}\mathbf{x}_{n+1}$. Knowing \mathbf{a}_{n+1} and $\mathbf{a}_n \bmod q$ (from the oracle) we can construct the $\mathbf{a}_i = q\mathbf{a}_{i+1} + (\mathbf{a}_i \bmod q)$ in reverse order, and with them all the \mathbf{x}_i up to $\mathbf{x}_1 = \mathbf{x}$. \square

Proof of Proposition 3.9

The result follows almost immediately from the three lemmas above. First, from Lemma 3.11, we have that, for any $\beta \leq \alpha$, we can construct an oracle W for $\text{SLWE}_{q, \phi_\beta}$ from our oracle for $\text{SLWE}_{q, \phi_\alpha}$, even for unknown β .

Second, Lemma 3.12 tells us that, for $\beta = \sqrt{(r\|\mathbf{e}\|/q)^2 + (rd/q)^2}$, we can solve q -BDD $_{\mathcal{L}^*,d}$ from our $D_{\mathcal{L},r}$ samples using W . The condition of $\beta \leq \alpha$ is satisfied for $d = \alpha q/(\sqrt{2}r)$. Indeed, plugging this value in and using that $\|\mathbf{e}\| \leq d$ yields

$$\beta = \sqrt{(r\|\mathbf{e}\|/q)^2 + (rd/q)^2} \leq \sqrt{\alpha^2/2 + \alpha^2/2} = \alpha.$$

Finally, since we can solve q -BDD $_{\mathcal{L}^*,\alpha q/(\sqrt{2}r)}$, we can also solve BDD $_{\mathcal{L}^*,\alpha q/(\sqrt{2}r)}$ by Lemma 3.13.

3.3 Finding an Oracle's Center

In the previous section, we presented a reduction from lattice problems to the search version of LWE. A natural question is whether one can make the reduction to the decision version of LWE directly, without relying on the decision-search equivalence from Lemma 3.4. This is achieved in [PRSD17]³, who introduce a methodology to find the *center* of a decision oracle under certain assumptions. Using this method, they replace the search oracle by a decision oracle in the reduction from BDD $_{\mathcal{L},d}$ to SLWE in Proposition 3.9.

Overall, there is little change in LWE due to Lemma 3.4⁴, but it will make a difference for Ring-LWE as we do not have a similar equivalence between the search and decision problems. Hence, the main motivation for this result does not come from LWE but rather from Ring-LWE.

In this section, we present the method for finding the center of a decision oracle, following the original approach in [PRSD17]. These results will be applied to Ring-LWE in Chapter 4. Besides, we present the main ideas of the reduction from BDD $_{\mathcal{L},d}$ to (decision) LWE. This reduction follows the same approach than the proof in Chapter 4, and we believe that reading Section 3.3.1 will help developing the intuition for it.

Motivation Recall the classical part in the previous section where a BDD instance $\mathbf{y} = \mathbf{x} + \mathbf{e}$ is solved using a SLWE oracle. Basically, we create LWE samples from the vector \mathbf{y} and from $D_{\mathcal{L},r}$ samples, where the secret \mathbf{s} is related to the lattice point \mathbf{x} that we are trying to find. If we replace our oracle by a decision LWE oracle, is this approach useful? *A priori* no, because our oracle will only confirm us that our samples come from an LWE distribution, and this will not really help us find \mathbf{s} .

However, let us see what happens if we slightly perturb the input vector $\mathbf{y}' = \mathbf{y} + \mathbf{z}$ used to generate the LWE samples. These will no longer be distributed exactly as before, but they will be close - hence, the probability that our oracle answers YES will vary slightly. Moreover, it is possible to prove that the probability distribution of the oracle has a central symmetry around the point $\mathbf{y}' = \mathbf{x}$, that we call *center*. Therefore, if we manage to find such a center by testing the probability distribution of the oracle on several points, we will be able to recover \mathbf{x} .

Oracles We will now introduce some notation. Let \mathcal{O} be a possibly randomized YES/NO oracle, i.e., a function $\mathcal{O} : S \rightarrow \{0, 1\}$ for a set S of possible queries. Oracles of this kind are often useful in reductions involving decision and promise problems. Let $p_{\mathcal{O}}(t) = \Pr[\mathcal{O}(t) = 1]$ (or simply $p(t)$) be the probability that \mathcal{O} answers YES on input t . For the univariate real case $S = \mathbb{R}$, we denote by $\mathcal{O}_s : \mathbb{R}^+ \rightarrow \{0, 1\}$ to the shifted oracle $\mathcal{O}_s(t) = \mathcal{O}(s + t)$.

³The full, 2020-revised version of the article (<https://eprint.iacr.org/2017/258>) is more detailed and easier to follow than the publication in STOC'17.

⁴In fact, the reduction does yield an improvement for LWE, as we no longer require q to be a prime number. This supports the choice of non-prime modulus in the Frodo-KEM cryptosystem and other LWE applications.

We will also deal with multivariate real oracles $\mathcal{O} : \mathbb{R}^k \times \mathbb{R}^+ \rightarrow \{0, 1\}$. In this case, a shifted oracle must fix the first input \mathbf{x} , such that $\mathcal{O}_s^*(t) = \mathcal{O}(\mathbf{x}, s + t)$. Our goal is to develop an algorithm for finding the center of multivariate oracles whose probability distributions have a central symmetry (i.e., the acceptance probability only depends on the euclidean distance to a centre \mathbf{z}^*). As an intermediate step, we will introduce a method for comparing two shifted univariate real oracles. We omit some technical details of the proofs as they do not provide any insight on LWE, but we present the main ideas.

Definition 3.14 (Oracle Comparison Problem). *Let $\epsilon \geq 0$, $r > 0$. Given an oracle $\mathcal{O} : \mathbb{R} \rightarrow \{0, 1\}$ and access to two shifted oracles $\mathcal{O}_{s_1}, \mathcal{O}_{s_2}$ for $s_1, s_2 \in [-r, r]$, the promise problem (ϵ, r) -OCP is to output YES if $s_2 \leq s_1 - \epsilon$ and NO if $s_2 > s_1$. Otherwise, any output is valid.*

Proposition 3.15. *Let \mathcal{O} be an oracle, $p(t)$ its acceptance probability, $\kappa \geq 200$, and s_1, s_2 two shifts of an instance of OCP that satisfy the following conditions: there exists a $p_\infty \in [0, 1]$ and a $t^* \geq s_1$ such that*

1. $p(t^*) - p_\infty \geq 1/\kappa$
2. $|p(t) - p_\infty| \leq 2e^{-t/\kappa}$ for all t
3. $p(t)$ is κ -Lipschitz⁵.

Then, there is a polynomial-time algorithm (in κ) that, on input κ , solves $(1/\kappa, \kappa)$ -OCP except with probability $e^{-\kappa}$. Each of the calls of the algorithm to the oracle is of the form $\mathcal{O}_{s_j}(i\Delta)$ for some $\Delta = \Delta(\kappa) < 1$, $0 \leq i \leq \text{poly}(\kappa)$ and $j \in \{1, 2\}$.

Before sketching the proof, let us develop the intuition behind these conditions. Even if this is not completely accurate, it might help to think of \mathcal{O} as a decision LWE oracle for a fixed distribution A_{s, ϕ_β} , where the input t represents the noise of the samples given to \mathcal{O} . In this setting, p_∞ is the probability of acceptance when the noise tends to infinity, where the samples become uniform. This fact is actually captured by condition (2), which basically says that $p(t)$ converges to p_∞ . Condition (1) means that there is a point t^* in which $p(t^*)$ is significantly larger than p_∞ ; this simply captures that the oracle has some advantage in distinguishing LWE samples from uniform samples. Condition (3) means that the acceptance probability is regular. Small changes in the error distribution will yield samples whose distributions are statistically close (see the start of Chapter 2). As the statistical distance cannot be amplified by any function, p cannot suffer large changes.

Proof Sketch. Consider the function $h(s) = \max_{t>0}(1+t)|p(s+t) - p_\infty|$. First, $h(s)$ is a decreasing function by definition. To see this, suppose $s_2 > s_1$. If the peak on $p(s+t)$ that determines $h(s_2)$ is reached at $t = t_2$, that maximum will appear on a larger $t_1 > t_2$ when $s = s_1$. Hence, the function will be larger on s_1 due to the factor $(1+t)$. In summary, any peak on s_2 will translate into a larger peak in s_1 , hence $h(s_1) \geq h(s_2)$. Therefore, in the NO instance of OCP, we will have $h(s_2) - h(s_1) \leq 0$. In the YES instance when $s_2 \leq s_1 - 1/\kappa$, one can prove that $h(s_2) - h(s_1) \geq 1/\text{poly}(\kappa)$ using conditions (1) and (2).

Hence, an algorithm that approximates h precisely on s_1, s_2 can be used to solve OCP. The algorithm is the following. On input $\kappa \geq 200$, and for each $i = 0, \dots, T$, make N calls to $\mathcal{O}_s(i\Delta)$. Then, set p'_i to the observed probability of the oracle at step i , and output $h' = \max_{0 \leq i \leq T}(1+i\Delta)|p'_i - p'_T|$. To prove that h' is a good approximation to h on input s , it suffices to see that $|h - h'| \leq 1/\text{poly}(\kappa)$. This is based on three observations that we sketch.

⁵A real-valued function $f : (a, b) \rightarrow \mathbb{R}$ is κ -Lipschitz if for any $x, y \in (a, b)$, $|f(x) - f(y)| \leq \kappa|x - y|$.

- Using the Chernoff bound, one can prove that $|p'_i - p(s+i\Delta)| \leq \kappa\Delta/10$ with probability larger than $1 - e^{-\kappa}$.
- By the Lipschitz property (condition (3)), we have that the discrete sampling is sufficient to approximate the actual function h - namely, $|p(t + \chi\Delta) - p(t)| \leq \kappa\Delta$ for any t and any $\chi \in [0, 1]$.
- By condition (2), we have that $p(s+T\Delta)$ approximates p_∞ , namely $|p(s+T\Delta) - p_\infty| \leq 2e^{-(s+T\Delta)/\kappa} \leq \kappa\Delta/10$.

Combining the three equations and recalling the definitions of the functions h, h' , the result follows. \square

Definition 3.16 (Oracle Hidden Center Problem). *Let $0 \leq \epsilon, \delta < 1$ and $\beta \geq 1$. The approximate search problem $(\epsilon, \delta, \beta)$ -OHCP is defined as follows. An instance is a scale parameter $d > 0$ and a multivariate oracle $\mathcal{O} : \mathbb{R}^k \times \mathbb{R}^+ \rightarrow \{0, 1\}$ whose (unknown) acceptance probability function p on input (\mathbf{z}, t) for $\|\mathbf{z} - \mathbf{z}^*\| \leq \beta d$ depends only on $t + \log\|\mathbf{z} - \mathbf{z}^*\|$, for an unknown hidden center $\delta d \leq \mathbf{z}^* \leq d$. The goal is to find a $\mathbf{z} \in \mathbb{R}^k$ such that $\|\mathbf{z} - \mathbf{z}^*\| \leq \epsilon d$.*

Notice that, in the previous definition, the probability function $p(t + \log\|\mathbf{z} - \mathbf{z}^*\|)$, behaves as a radial function when t is fixed - hence the notion of a hidden center. To solve OHCP, the algorithm resembles some optimization algorithms (such as the interior-point method), in the sense that it will iterate approaching the center \mathbf{z}^* .

Theorem 3.17. *Let $\mathcal{O} : \mathbb{R}^k \times \mathbb{R}^+ \rightarrow \{0, 1\}$ be a multivariate oracle, $\kappa \geq 20 \log(k+1)$, and $p(t)$ the acceptance probability of \mathcal{O} on input $\mathcal{O}(\mathbf{0}, t)$ of an instance of OHCP that satisfies the following conditions: there exists a $p_\infty \in [0, 1]$ and a $t^* \geq 0$ such that*

1. $p(t^*) - p_\infty \geq 1/\kappa$,
2. $|p(t) - p_\infty| \leq 2e^{-t/\kappa}$ for all $t \in \mathbb{R}^+$,
3. $p(t)$ is κ -Lipschitz.

Then, there is a polynomial-time algorithm (in κ, k) that, on input κ , solves $(e^{-\kappa}, e^{-\kappa}, 1+1/\kappa)$ -OHCP except with probability $e^{-\kappa}$. Each of the calls of the algorithm to the oracle is of the form $\mathcal{O}(\cdot, i\Delta)$ for some $\Delta = \Delta(\kappa) < 1$, $0 \leq i \leq \text{poly}(\kappa, k)$.

Proof Sketch. Without loss of generality, we can take $d = 1$ as the problem is invariant under scaling. Overall, the algorithm constructs a sequence $\{\mathbf{z}_i\}_i$ that approaches the center of the oracle \mathbf{z}^* , such that $\|\mathbf{z}_{i+1} - \mathbf{z}^*\| \leq \|\mathbf{z}_i - \mathbf{z}^*\|$. We separate the proof sketch in three components for clarity. First, we will use Proposition 3.15 to build an auxiliary algorithm that is able to guide a sequence towards the center. Second, we will build the algorithm that constructs that sequence. Third, we prove the correctness of the algorithm, i.e., that the sequence approaches the center as desired, and that at the end we have a point sufficiently close (depending on κ) to the center.

1. **GUIDANCE.** We aim for a distinguisher that, given two points $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^k$ such that $e^{-\lambda} \leq \|\mathbf{x}_j - \mathbf{z}^*\| \leq 1 + 1/k$, tells us which point is closer to \mathbf{z}^* . Let $\mathcal{O}^*(t) = \mathcal{O}(\mathbf{0}, t)$, $\mathcal{O}_1(t) = \mathcal{O}(\mathbf{x}_1, t)$, $\mathcal{O}_2(t) = \mathcal{O}(\mathbf{x}_2, t)$. Notice that the latter two correspond to a shift of \mathcal{O}^* . Indeed, if we let $s_i = \log\|\mathbf{x}_i - \mathbf{z}^*\|$, we have that

$$\mathcal{O}_{s_i - \log\|\mathbf{z}^*\|}^*(t) = \mathcal{O}(\mathbf{0}, t + \log\|\mathbf{x}_i - \mathbf{z}^*\| - \log\|\mathbf{z}^*\|),$$

and its probability distribution (by Definition 3.16) is

$$p((t + \log\|\mathbf{x}_i - \mathbf{z}^*\| - \log\|\mathbf{z}^*\|) + \|\mathbf{0} - \mathbf{z}^*\|) = p(t + \log\|\mathbf{x}_i - \mathbf{z}^*\|),$$

which is exactly the probability distribution of $\mathcal{O}_i(t)$. Hence, after checking that the conditions on the oracle hold, we can apply Proposition 3.15 to obtain an algorithm that distinguishes the cases $s_1 \leq s_2$ and $s_2 \leq s_1 - 1/\lambda$ with probability at most e^λ . Equivalently, our distinguisher works for the cases $\|\mathbf{x}_1 - \mathbf{z}^*\| \leq \|\mathbf{x}_2 - \mathbf{z}^*\|$ and $\log\|\mathbf{x}_2 - \mathbf{z}^*\| \leq \log\|\mathbf{x}_1 - \mathbf{z}^*\| - 1/\lambda$.

2. **BUILDING THE SEQUENCE.** Let $\mathbf{z}_0 = \mathbf{0}$. For $i = 0, \dots, T$, we will sample $j \in \{1, \dots, k\}$, $x \in [0, 1]$ and $\sigma \in \{\pm 1\}$ uniformly at random, and set $\mathbf{v}_i = \sigma e^{-2\kappa x} \mathbf{e}_j / \sqrt{\kappa^2 k}$. Candidates \mathbf{x}_i for \mathbf{z}_i will be chosen as $\mathbf{x}_{i+1} = \mathbf{z}_i + \mathbf{v}_i$. Then, we set $\mathbf{z}_{i+1} = \mathbf{x}_{i+1}$ if $\log\|\mathbf{x}_{i+1} - \mathbf{z}^*\| \leq \log\|\mathbf{z}_i - \mathbf{z}^*\| - 1/\lambda$, and $\mathbf{z}_{i+1} = \mathbf{z}_i$ otherwise (i.e., we do not advance the sequence). To decide which is the case, we use the previous distinguisher.

3. **CORRECTNESS.** For correctness, we need to prove: (a) that the guidance algorithm answers correctly for all steps i except with small probability, and (b) that $s_{T+1} = \log\|\mathbf{z}_{T+1} - \mathbf{z}^*\| \leq -\kappa$, also except with small probability.

For (a), we proceed by induction on i . In particular, for every step, we need to ensure that the parameters are in the allowed ranges for the oracle to work. In particular, that $s_i \in [-\lambda, \log(1 + 1/\kappa)]$ to ensure that the assumption on (1) holds. For (b), the argument is mostly probabilistic, and consists in ensuring that sufficiently many choices of \mathbf{v}_i make progress in the sequence. Notice that we are approaching a center, hence there must exist at least one coordinate in which significant progress can be made. \square

3.3.1 Application to LWE

As mentioned previously, the classical part of the IS presented in Proposition 3.9 can be improved using Theorem 3.17 by using an oracle for the decision version LWE instead of search SLWE. We present it below; as we are only interested in the intuition, we will skip several details in the proof.

Proposition 3.18 (Classical part of the IS, [PRSD17]). *Let $\epsilon = \epsilon(n) > 0$ be negligible, $q \geq 2$, $\alpha \in (0, 1)$. Given access to an oracle that solves $\text{LWE}_{q, \phi_\alpha}$ given a polynomial number of samples, there exists an efficient algorithm that, for any n -dimensional lattice \mathcal{L} , $r > \sqrt{2}q\eta_\epsilon(\mathcal{L})$, solves $\text{BDD}_{\mathcal{L}^*, \alpha q / (\sqrt{2}r)}$ using a polynomial number of samples from different $D_{\mathcal{L}, r_i}$, where $r_i \geq r$.*

If we look back to our Iterative Step, our reduction now requests samples from different spherical Gaussians $D_{\mathcal{L}, r_i}$, where $r_i \geq r$, so the IS must be able to provide such samples. The original quantum step in [Reg09] does not explicitly mention this. Nevertheless, the proof works in that case too, so the generalized version of Proposition 3.10 does consider it. Besides, in the initial bootstrapping, sampling from wider Gaussian distributions is also permitted. To improve readability, we recall the main lemma in the proof of the classical step.

Lemma 3.12. *Let $\text{BDD}_{\mathcal{L}^*, d}$ be instance given by a vector $\mathbf{y} = \mathbf{x} + \mathbf{e}$ such that $\mathbf{x} \in \mathcal{L}^*$ and $\|\mathbf{e}\| \leq d$. There is a probabilistic algorithm that, given such instance, an $r \geq \sqrt{2}q\eta(\mathcal{L})$, and polynomially many samples from $D_{\mathcal{L}, r}$, outputs samples from A_{s, ϕ_β} where $\beta = \sqrt{(r\|\mathbf{e}\|/q)^2 + (rd/q)^2}$ and $\mathbf{s} = (\mathcal{L}^*)^{-1}\mathbf{x} \bmod q$.*

Proof Sketch. (Proposition 3.18). Let $\kappa = \text{poly}(n)$. We start with a $\text{BDD}_{\mathcal{L}^*, d}$ instance as in Lemma 3.12, where \mathbf{x} is our target vector and $\mathbf{y} = \mathbf{x} + \mathbf{e}$ is known. Using the oracle for $\text{LWE}_{q, \phi_\alpha}$, we want to create an oracle $\mathcal{O}(\mathbf{z}, t)$ compatible with Definition 3.16, that takes as

input a shift vector \mathbf{z} and a scale parameter t . The probability of the oracle will depend only on $e^t \|\mathbf{z} - \mathbf{e}\|$; hence, the oracle will have in \mathbf{e} a hidden center. After ensuring that the conditions are met, we will apply Theorem 3.17 to find \mathbf{e} .

A call $\mathcal{O}(\mathbf{z}, t)$ proceeds as follows. First, it takes samples from $D_{\mathcal{L}, e^t r}$ and applies Lemma 3.12 on the vector $\mathbf{y} - \mathbf{z}$, radius r and bound d to obtain a set of samples $M_{\mathbf{z}, t}$. Finally, \mathcal{O} calls the $\text{LWE}_{q, \phi_\alpha}$ oracle on input $M_{\mathbf{z}, t}$ and returns its answer. Notice that calls to $\mathcal{O}(\mathbf{z}, t)$ with small \mathbf{z} will result in samples generated from points very close to \mathbf{y} . On the other hand, if t is large, the distribution will present a large error, eventually becoming uniform.

The correctness of the procedure follows from the choice of the tolerance parameter κ . If κ is sufficiently small, the procedure in Theorem 3.17 will output a value $\tilde{\mathbf{e}}$ which is close enough to the center, say $\|\mathbf{e} - \tilde{\mathbf{e}}\| \leq 2^{-n} \lambda_1(\mathcal{L}^*)$. To find the exact vector \mathbf{e} (or, equivalently, \mathbf{x}), we can run Babai's nearest plane algorithm (Lemma 2.23).

Hence, it suffices to prove that the oracle \mathcal{O} is a valid instance of the OCP problem and that it satisfies the regularity conditions imposed by Theorem 3.17. The first part follows from careful bounding of the parameters and from the fact that the probability of \mathcal{O} depends only on $e^t \|\mathbf{z} - \mathbf{e}\|$, as the distribution of the samples $M_{\mathbf{z}, t}$ depends only on that value. Regularity can be proven from the observation that, when $t \rightarrow \infty$, the distribution of $M_{\mathbf{z}, t}$ is uniform (the noise hides all the LWE structure), giving us the desired behaviour for p_∞ . The Lipschitz condition holds as for any $t_1 \approx_\epsilon t_2$, the statistical distance of the distributions of $M_{\mathbf{z}, t_1}$ and $M_{\mathbf{z}, t_2}$ is very small (Lemma 2.7). Therefore, the acceptance probability of the oracle cannot vary a lot among both distributions. \square

Chapter 4

Ring-LWE

The Ring Learning with Errors problem is a variant of LWE introduced in [LPR10]. Previous work, such as [SSTX09] or even the NTRU cryptosystem [HPS98], already incorporated the main idea of RLWE for rings of the form $\mathbb{Z}[x]/(x^n \pm 1)$. Recently, RLWE has been used on many cryptographic primitives, and its security, based on the hardness of computational problems on ideal lattices, has been gradually strengthened. In this chapter, we first introduce the RLWE problem, and then we present the strongest hardness results available in the literature.

Our exposition follows [LPR10, PRSD17] and builds on the ideas in Chapter 3. The reductions from ideal lattice problems to Ring-LWE follow the same structure; we will have an iterative step with classical and quantum parts, where the classical part consists in a reduction from R -GDP directly to decision-RLWE. The chapter concludes with a discussion on related problems and a survey of variants of Ring-LWE.

4.1 Overview

We start by providing the definition of the RLWE distribution, as we did for LWE in Chapter 3. The idea is very similar; we simply replace the vectors of integers by elements from the ring of integers $R = \mathcal{O}_K$ of a number field K . Following the notation from Section 2.2, let $\mathbb{T} = K_{\mathbb{R}}/R^{\vee}$ and denote $\mathcal{I}_q = \mathcal{I}/q\mathcal{I}$ for any fractional ideal \mathcal{I} of R and a rational integer $q \geq 2$. The space \mathbb{T} can be thought of as the continuous fundamental domain of the ideal lattice generated by the codifferent ideal R^{\vee} .

Definition 4.1 (RLWE distribution). *For $s \in R_q^{\vee}$ and for an error distribution ψ over $K_{\mathbb{R}}$, a sample from the RLWE distribution $A_{s,\psi}$ over $R_q \times \mathbb{T}$ is obtained by sampling an $a \in R_q$ uniformly at random, sampling $e \in K_{\mathbb{R}}$ from ψ , and outputting the pair $(a, as/q + e \bmod R^{\vee})$.*

There is a significant efficiency gain in using Ring-LWE as opposed to LWE. For a single LWE sample $(\mathbf{a}, b = \langle \mathbf{a}, \mathbf{s} \rangle / q + e)$ we required a fresh vector $\mathbf{a} \in \mathbb{Z}_q^n$ and we obtained one pseudorandom value $b \in \mathbb{R}/\mathbb{Z}$. On the other hand, a single Ring-LWE sample $(a, b = as/q + e \bmod R^{\vee})$ requires a ring element a (which is somewhat equivalent to a vector of \mathbb{Z}_q^n as the ring has degree n), but yields a full vector $b \in \mathbb{T}$ of n coordinates. If we assume the hardness of the decision-RLWE problem that we introduce below, we will get n pseudorandom values from one sample.

In some works such as [RSW18], the distribution is slightly different as the error is directly sampled from an elliptical Gaussian distribution $D_{\mathbf{r}}$ (besides, the definitions for \mathbb{T} may vary). In practice, this makes no difference as ψ will be an elliptical Gaussian distribution in all our results. The RLWE problem can be stated both on its search and decision variants.

Definition 4.2 (Search-RLWE). *For a family Ψ of distributions over $K_{\mathbb{R}}$, the problem Search-RLWE, denoted by $\text{SRLWE}_{q,\Psi}$, is defined as follows: given arbitrarily many independent samples from the RLWE distribution $A_{s,\psi}$ for some fixed secret $s \in R_q^\vee$ and $\psi \in \Psi$, find s .*

Definition 4.3 (Average-case Decision-RLWE). *For a distribution Γ over a family of error distributions over $K_{\mathbb{R}}$, the average-case Decision-RLWE problem, denoted by $\text{RLWE}_{q,\Gamma}$, is the following: with non-negligible probability, distinguish between arbitrarily many independent samples from the RLWE distribution $A_{s,\psi}$ for a random (uniform) choice of $s \in R_q^\vee$ and of $\psi \in \Gamma$, and the same number of uniformly random and independent samples from $R_q \times \mathbb{T}$.*

Error distributions The previous definitions might be somewhat obscure, since we require a family of error distributions Ψ (and a distribution over a family of error distributions Γ). We can think of Ψ as a family of elliptical Gaussians $D_{\mathbf{r}}$ over $K_{\mathbb{R}}$, as introduced in Chapter 2. We are particularly interested in families of the form Ψ_α in which the radius is bounded, $\mathbf{r} \leq \alpha$ (meaning that $r_i \leq \alpha$ for each i), for every $D_{\mathbf{r}} \in \Psi_\alpha$. The reason why we work with elliptical Gaussians, and not spherical as before, is that the image of an element under the canonical embedding does not yield spherical distributions. For instance, recall the reduction from BDD to SLWE in Proposition 3.9, and consider what happens if we have a R -BDD instance instead (where $y = x + e$ and all elements live in $\mathcal{I} \subset R$). In this case, we need to project e into $H \cong K_{\mathbb{R}}$ to convert it to the noise of Ring-LWE samples, and we will obtain a vector $(\sigma_1(e), \dots, \sigma_n(e)) \in H$. The distribution of such a vector depends on each of the σ_i , hence it will not be spherical but elliptical, where each coordinate is bounded by $\|e\|_\infty = \max_i |\sigma_i(e)|$. This example will become clearer in Section 4.2, when we introduce the RLWE sample generation from R -BDD instances.

Definition 4.4 (Family Ψ_α). *For $\alpha > 0$, the family Ψ_α is the (uncountable) set of all elliptical Gaussian distributions $D_{\mathbf{r}}$ over $K_{\mathbb{R}}$ in which $\mathbf{r} \leq \alpha$.*

When we define a distribution Γ over a family of elliptical Gaussians, we actually have a distribution over different radius vectors \mathbf{r} . Obtaining one sample from Γ is equivalent to sampling \mathbf{r} according to the distribution, and outputting $D_{\mathbf{r}}$. The following is the distribution that we consider for the decision problem, as in [PRSD17]. Recall that we defined $G = \{\mathbf{r} \in (\mathbb{R}^+)^n : r_{i+s_1+s_2} = r_{i+s_1}, i = 1, \dots, s_2\}$ as the set of possible radial vectors for an elliptical Gaussian distribution when we work with the space H , where ideal lattices live.

Definition 4.5 (Distribution Γ_α). *Let $\alpha > 0$ and $f(n) = \omega(\sqrt{\log n})$ be a fixed that grows asymptotically faster than a $\sqrt{\log n}$. A sample from Γ_α is an elliptical Gaussian $D_{\mathbf{r}}$ such that $\mathbf{r} \in G$ is sampled as:*

$$\begin{cases} r_i = \alpha \sqrt{(x_i^2 + f(n)^2)/2} & \text{for } 1 \leq i \leq s_1 & \text{where } x_i \leftarrow D_1 \\ r_i = r_{i+s_2} = \alpha \sqrt{(x_i^2 + y_i^2 + f(n)^2)/2} & \text{for } s_1 \leq i \leq s_1 + s_2 & \text{where } x_i, y_i \leftarrow D_{1/\sqrt{2}} \end{cases}$$

Sampling from Γ_α yields error distributions whose radius are of size $O(\alpha \cdot \omega(\sqrt{\log n}))$, essentially $O(\alpha \cdot \log n)$. The choice of Γ_α will become clearer in Section 4.2.

Design choices and primal RLWE Before proceeding with the hardness results, we provide further intuition as to why is RLWE defined as above. First, RLWE is meant to be practical and implementable, so it may seem strange that the second component of the RLWE distribution $A_{s,\psi}$ belongs to the continuous domain \mathbb{T} . In practice, the samples can be discretized, as in the case of Plain-LWE. Namely, as proven in [LPR13] (Lemma 2.23), the hardness of Decision Ring-LWE is preserved if the error distributions are suitably discretized.

Second, notice that the error e in Definition 4.1 is sampled from a Gaussian distribution ψ and then reduced modulo R^\vee . Intuitively, RLWE should be harder when such errors are large, and this is indeed the case. However, if the error is too large, the distribution becomes statistically indistinguishable from uniform and hence RLWE becomes impossible to solve. This occurs precisely when the size of e (roughly the width of ψ , by Lemma 2.13) exceeds the smoothing parameter of R^\vee , as presented in Lemma 2.9.

Third, one might question why is the dual ideal R^\vee used instead of the primal R , which seems like the natural choice. The answer is that both choices are equivalent, as we prove below. Nevertheless, R^\vee is more convenient for our proofs, since it will arise naturally in the R -BDD to SRLWE reduction.

Definition 4.6 (Primal RLWE distribution). *Let $\mathbb{T}^P = K_{\mathbb{R}}/R$. For $s \in R_q$ and for an error distribution ψ over $K_{\mathbb{R}}$, a sample from the primal RLWE distribution $A_{s,\psi}^P$ over $R_q \times \mathbb{T}^P$ is obtained by sampling an $a \in R_q$ uniformly at random, sampling $e \in K_{\mathbb{R}}$ from ψ , and outputting the pair $(a, as/q + e \bmod R)$.*

Proposition 4.7 (Equivalence of dual and primal RLWE [RSW18]). *Define the primal (average-case) Decision-RLWE problem as in Definition 4.3, but sampling from $A_{s,\psi}^P$ (Definition 4.6). If the error distributions are elliptical Gaussians, both problems are equivalent but for a scaling factor on the error.*

Proof. First, notice that there is a natural reduction from primal to dual given by the inclusion $f : R \rightarrow R^\vee$. Indeed, f maps $A_{s,\psi}^P$ to $A_{s,\psi}$ and the uniform distribution on $R_q \times \mathbb{T}^P$ to the uniform distribution on $R_q \times \mathbb{T}$.

For the dual to primal reduction, let $t \in (R^\vee)^{-1}$ such that tR^\vee and qR are coprime, as in Lemma 2.44. By Lemma 2.45, t induces a bijection $\theta_t : R^\vee/(qR^\vee) \rightarrow R^\vee(R^\vee)^{-1}/qR^\vee(R^\vee)^{-1} = R_q$ given by $\theta_t(x) = tx$.

Now let (a, b) be a sample from A_{s,D_r} . We are interested in the distribution of $(a, \theta_t(b))$. First, we have that $\theta_t(b) = \theta_t(as/q + e) = a(ts)/q + te$, where $ts, te \in R_q$. The product te follows a distribution $D_{r'}$ (recall that the product of ring elements results in coordinate-wise multiplication under σ), where $r'_i = r_i|\sigma_i(t)|$. Hence, $(a, \theta_t(b))$ is distributed as $A_{ts,D_{r'}}$.

On the other hand, if (a, b) is uniformly distributed, then $\theta_t(b) = tb$ is uniformly distributed in \mathbb{T}^P independently of a , since θ_t is an isomorphism. We conclude that if (a, b) is a sample for dual Decision-RLWE, $(a, \theta_t(b))$ is a valid sample for primal Decision-RLWE. \square

The equivalence is, however, not total. The proof works for arbitrary rings, but does not control the transformation of the error distribution under the reduction. When we consider a primal instance, the error distribution is scaled with respect to the dual distribution (by the factor $\sigma_i(t)$). The hardness results for Ring-LWE that we introduce only hold for the dual variant. Therefore, one must be very careful when using primal RLWE in a cryptographic construction, as the resulting error distribution in the dual might be too narrow to provide the desired level of security. In general, error control is a complex and well-studied problem which is also useful in the study of variants such as Polynomial LWE [RSW18]. In Section 4.3, we will expand this discussion and survey some attacks based on this weakness.

Initially, the only setting in which primal RLWE was considered is the case $R = \mathbb{Z}[x]/(x^n + 1)$ for $n = 2^k$. The reason is that the dual ideal R^\vee is a scaled isometry of R . Therefore, all results extend immediately [LPR10]. Rings of such form are particularly interesting as $x^n + 1$ is a cyclotomic polynomial with some nice properties (for instance, bounded factors for noise control and fast ring arithmetic), which makes them widely used in practice. This equivalence was later extended to arbitrary cyclotomic polynomials [DD12]. Furthermore, security in the case of $R = \mathbb{Z}[x]/(x^n + 1)$ was proven in an early work by Stehlé et. al.,

using the so-called polynomial embedding, which consists on considering the cyclic lattice generated by the coefficients of the polynomials in R [SSTX09]. We will also extend this discussion in Section 4.3.

4.2 Hardness

In this section, we prove the main hardness result for Ring-LWE, which consists in a quantum reduction from ideal lattice problems to the decision form of RLWE. As in the case of LWE, we have two versions of the result. The first (Theorem 4.8) is a generalized result which reduces R -SIVP $_\gamma$ to RLWE for linear approximation factors. The second (Theorem 4.9), which is the result we prove, is a reduction from R -DGS $_\varphi$ to RLWE. To obtain Theorem 4.8 from Theorem 4.9, we need the standard reductions mentioned in Section 2.1.3. In contrast to Theorem 3.6 for LWE, the R -GapSVP $_\gamma$ problem is not introduced since this problem is easy on ideal lattices.

The factor $\omega(1)$ denotes a fixed function that grows asymptotically faster than a constant, and plays the role of an approximation factor. Notice that we already introduced an $\omega(\cdot)$ function in Definition 4.5. We will also use $o(1)$, which represents a fixed function whose limit tends to zero, and can be seen as the inverse of $\omega(1)$.

Theorem 4.8 (Main theorem, extended). *Let K be a number field of degree n and $R = \mathcal{O}_K$ its ring of integers. Let $q \geq 2$, $\alpha \in (0, 1)$ such that $\alpha q > \omega(1)$. If there is an efficient algorithm that solves $\text{RLWE}_{q, \Gamma_\alpha}$, then there exists an efficient quantum algorithm that solves R -SIVP $_\gamma$ for $\gamma = \max\{\tilde{\omega}(\sqrt{n}/\alpha), O(n)\}$.*

Theorem 4.9 (Main theorem). *Let K be a number field of degree n and $R = \mathcal{O}_K$ its ring of integers. Let $q \geq 2$, $\alpha \in (0, 1)$ such that $\alpha q > 2\omega(1)$. There is a polynomial time quantum reduction from R -DGS $_\gamma$ (for a negligible $\epsilon = \epsilon(n) > 0$) to $\text{RLWE}_{q, \Gamma_\alpha}$, for any*

$$\gamma = \max \left\{ \sqrt{2}\eta_\epsilon(\mathcal{I})\omega(1)/\alpha, \sqrt{2n}/\lambda_1(\mathcal{I}^\vee) \right\}.$$

In [PRSD17], they use the inequality $\eta_\epsilon(\mathcal{I}) > \omega(\sqrt{\log n})/\lambda_1(\mathcal{I}^\vee)$ to show that the first term $\gamma = \sqrt{2}\eta_\epsilon(\mathcal{I})\omega(1)/\alpha$ dominates in Theorem 4.9, as long as $\alpha < \sqrt{\log n/n}$ which is the usual case in applications. The reduction becomes tighter (i.e., works for a smaller γ) when the smoothing parameter of \mathcal{I} is large, meaning that \mathcal{I} is a sparse ideal lattice (and that \mathcal{I}^\vee is dense). A larger α implies a wider Ring-LWE error distribution, hence the problem is harder and γ becomes smaller as expected.

This section is structured in the following way. First, we present an overview of the reduction, highlighting the differences and similarities with LWE, besides the main drawbacks. Second, we introduce the main theorem and the Iterative Step for Ring-LWE. Third, we explore the classical step of the IS. Finally, we use the method for finding the center of the oracle (Section 3.3) to extend our reduction to decision-RLWE for any number field.

4.2.1 Overview

Approach. The proof of Theorem 4.9 follows the same approach as the reduction for LWE that we presented in Chapter 3. The core part of the reduction is an Iterative Step which again has two parts, a classical part and a quantum part (which remains the same). The main difference is that we now work with ring elements instead of integer vectors. To embed ring elements into (ideal) lattices, the canonical embedding is required.

As opposed to Chapter 3, where we first presented a reduction to search-LWE, here we reduce to decision-RLWE directly. We do so by including the techniques in Section 3.3 in our

proofs (finding an oracle’s center). Nevertheless, we will need a method to generate RLWE samples from a R -BDD instance that can be seen as a stand-alone classical part that uses a search-RLWE oracle. Again, the reductions are worst-case, meaning that we can transform any ideal lattice problem into a RLWE instance.

Differences with LWE. The most significant differences with to LWE appear in the classical step. The first thing to notice is that we now solve a R -GDP $_{\mathcal{I}^\vee}$ instance using the RLWE oracle, as opposed to BDD. Both problems are essentially the same, and if we can solve R -BDD $_{\mathcal{I}^\vee, d}$, we will be able to solve R -BDD $_{\mathcal{I}^\vee, g}$ for some $g = \xi(n) \cdot d$ (recall that the norm of vectors sampled from $D_{\mathcal{I}, g}$ will have norms bounded by d with very high probability). However, the main reason why we work with R -GDP is the self-reducibility result from Proposition 2.24, which ensures that if we can solve GDP $_{\mathcal{L}, g}$ with non-negligible probability for some g , we will be able to solve it with overwhelming probability for $g' = o(1)g$. This is also one of the reason why the functions $\omega(1)$ and $o(1)$ appear; we need to take into account approximation factors that partially vanish when we use the self-reducibility result.

The fact that ring elements are embedded via the canonical embedding also complicates some of the steps. For example, we need to work with the ℓ_∞ norm (in the R -BDD instances), since we need to bound all the coordinates of our embedded elements (in order to control our error distributions). Besides, we no longer have the direct correspondence between vectors and lattice vectors that we used in Lemma 3.12 to relate the secret and the samples. To circumvent this difficulty, we will make use of the efficient bijection from Lemma 2.44 and Lemma 2.45 to map $R_q \leftrightarrow \mathcal{I}_q$. In this way, we can establish a one-to-one correspondence between lattice (ideal) elements and ring elements, which form the Ring-LWE instance.

Another remark is that our reductions require samples from different elliptical Gaussians, i.e., from distributions $D_{\mathcal{I}, \mathbf{r}}$ for several different \mathbf{r} in each step. This feature is required in the search of the decision oracle’s center, where our samples need to be diverse. In this part, we replace the single oracle in the proof of Proposition 3.18 by $s_1 + s_2$ oracles, roughly one per embedding coordinate; we need to find the center of each of them (recall that $n = s_1 + 2s_2$).

Finally, one of the main differences is hidden in the error distributions. In the case of LWE, it suffices to have an oracle for a single error distribution ϕ_α due to the self-reducibility result in Lemma 3.11. On the other hand, our RLWE $_{q, \Gamma_\alpha}$ oracle is required to solve decision Ring-LWE for *any* elliptical Gaussian sampled from Γ_α , which may yield quite diverse Gaussians.

Drawbacks. The former fact significantly weakens the result, making it less satisfactory than the one for LWE. In an attempt to overcome this limitation, both [LPR10] and [PRSD17] present a reduction that holds for fixed spherical error distributions. Nevertheless, the error width increases by a factor of $O((mn)^{1/4})$, where m is the number of samples required by the oracle. The dependency on the number of samples can be significant in applications and differs to the results obtained for plain LWE. We will comment on this result in Section 4.2.4.

We remark that the results for Ring-LWE are intrinsically weaker than the results for LWE, since the reductions are only valid for ideal lattices (and only for the dual form of Ring-LWE where $s \in R_q^\vee$). Whether computational problems on ideal lattices are as hard as their counterparts in general lattices is not clear; we comment on this in Section 4.3.

4.2.2 The Iterative Step

As we mentioned, one of the main differences with the LWE reduction is that we require Ring-LWE samples from elliptical error distributions $D_{\mathcal{I}, \mathbf{r}}$, for different vectors \mathbf{r} . The set of vectors that we will need belong to the set $W_{r, \zeta, T}$, which is a finite subset of G of size

$(s_1 + s_2)(T + 1)$. The vectors in $W_{r,\zeta,T}$ are slight perturbations of the spherical vectors $\mathbf{r} = (r, r, \dots, r)$ on a single coordinate. Recall that s_1 is the number of real embeddings and s_2 is the number of pairs of complex embeddings.

Definition 4.10. *Let $r \geq 0$, $\zeta = 1/\text{poly}(n) > 0$, and $T = \text{poly}(n) \geq 1$. Let also $i = 1, \dots, s_1 + s_2$ and $j = 0, \dots, T$. We denote $W_{r,\zeta,T}$ to the set of vectors $\mathbf{r}_{i,j} \in G$ such that all coordinates are equal to r except the i -th coordinate (and the $(i + s_2)$ -th if $i > s_1$), which is equal to $r(1 + \zeta)^j$.*

The proof of Theorem 4.9 follows from the Iterative Step that we present below, which preserves the structure of the IS in Chapter 3, as summarized in Figure 4.1. Again, we have a classical part and a quantum part that are applied as before. The latter is unchanged, whereas the former is adapted to the ring setting. One of the main differences is that we use a decision oracle, achieving a reduction to decision RLWE. This is a consequence of the novel approach in [PRSD17], which follows the same argument that we presented in Section 3.3 with the modification of the classical part (Proposition 3.18). As we mentioned, other differences are the use elliptical discrete Gaussians such as $D_{\mathcal{I},\mathbf{r}}$, the increased complexity of the parameter set, and that the intermediate lattice problem is now GDP instead of BDD, since the offset will be sampled from a Gaussian distribution.

Proposition 4.11 (Iterative Step). *Let $\epsilon = \epsilon(n) > 0$ be negligible, $q \geq 2$, $\alpha \in (0, 1)$. Given access to an oracle that solves $\text{RLWE}_{q,\Gamma_\alpha}$ given a polynomial number of samples, there exists an efficient quantum algorithm that, for any fractional ideal $\mathcal{I} \subset K$, a radius $r > \sqrt{2}q\eta_\epsilon(\mathcal{I})$ such that $r' = r\omega(1)/(\sqrt{2}\alpha q) > \sqrt{2n}/\lambda_1(\mathcal{I}^\vee)$, a polynomial number of samples from $D_{\mathcal{I},\mathbf{r}}$ for each $\mathbf{r} \in W_{r,\zeta,T}$, and a vector $\mathbf{r}' \geq r'$, produces an independent sample of $D_{\mathcal{I},\mathbf{r}'}$.*

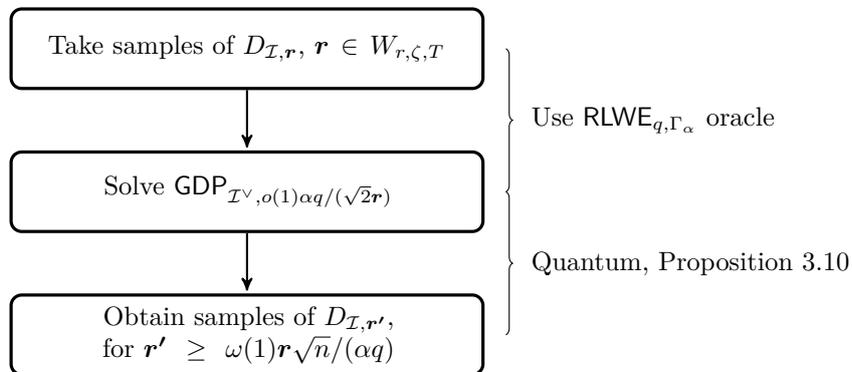


Figure 4.1: Diagram of the Iterative Step for Ring-LWE

In an attempt to clarify the role of the parameters in the IS, we describe all its inputs below.

- The parameters $\epsilon, q \geq 2, \alpha \in (0, 1)$. As in LWE, an instance of $\text{RLWE}_{q,\Gamma_\alpha}$ where α is small involves easier lattice problems, since we will have narrower error distributions. If we have a small q , our α is required to be larger. The parameter ϵ determines the tolerance of the smoothing parameter of the ideal lattices; a small ϵ ensures that the error distributions that we obtain in our reductions are tight. Also, notice that we do not need the factorization of q ; as the reduction is quantum and $q = \text{poly}(n)$, we can factor q in polynomial time if needed.

- Access to an oracle that solves $\text{RLWE}_{q,\Gamma_\alpha}$ given a polynomial number of samples. This is a *decision* oracle, which appears because we include the main result in [PRSD17]. On the other hand, we need an oracle for the distribution Γ_α defined over a family of error distributions, whereas for LWE it suffices to have an oracle for a single distribution ϕ_α .
- Any fractional ideal $\mathcal{I} \subset K$, meaning that the reduction works for any ideal lattice (i.e., we have a worst-case reduction).
- A radius $r > \sqrt{2q}\eta_\epsilon(\mathcal{I})$ such that $r' = r\omega(1)/(\sqrt{2}\alpha q) > \sqrt{2n}/\lambda_1(\mathcal{I}^\vee)$. The role of r is to characterize the width of the inputs samples from $D_{\mathcal{I},\mathbf{r}}$, since \mathbf{r} is a vector in $W_{r,\zeta,T}$ (recall that all coordinates except one of \mathbf{r} are equal to r , and $\mathbf{r} \geq r$). The lower bound on r guarantees that \mathbf{r} is sufficiently above the smoothing condition $\eta_\epsilon(\mathcal{I})$, which is required in Lemma 4.13 for handling the error distributions. It also imposes a bound on how narrow our distributions can become, and hence on the parameters of the R -DGS problem that we solve.

The value $\omega(1)/(\sqrt{2}\alpha q)$ is the largest width reduction factor that we can achieve for our samples in a single IS. The limit $r' > \sqrt{2n}/\lambda_1(\mathcal{I}^\vee)$ is imposed by the quantum part (Proposition 3.10).

- A vector $\mathbf{r}' \geq r'$ of our choice, that characterizes the output. We can request our algorithm to output samples from *any* distribution $D_{\mathcal{I},\mathbf{r}'}$, as long as each of its coordinates $r'_i \geq r'$.

Proposition 4.12 (Classical part of the IS). *Let $\epsilon = \epsilon(n) > 0$ be negligible, $q \geq 2$, $\alpha \in (0, 1)$. Given access to an oracle that solves $\text{RLWE}_{q,\Gamma_\alpha}$ given a polynomial number of samples, there exists an efficient algorithm that, given any fractional ideal $\mathcal{I} \subset K$, a radius $r > \sqrt{2q}\eta_\epsilon(\mathcal{I})$, and the factorization of q , solves $R\text{-GDP}_{\mathcal{I}^\vee, o(1)\alpha q/(\sqrt{2}r)}$ using a polynomial number of samples from $D_{\mathcal{I},\mathbf{r}}$ for each $\mathbf{r} \in W_{r,\zeta,T}$.*

Proposition 3.10 (Quantum part of the IS). *Given any n -dimensional lattice \mathcal{L} , a real $g < \lambda_1(\mathcal{L}^*)/(\sqrt{8n})$, a vector \mathbf{r} such that $r_i \geq 1$ for $i = 1, \dots, n$, and an oracle that solves $\text{GDP}_{\mathcal{L}^*,g}$ with overwhelming probability, there is an efficient quantum algorithm that outputs an independent sample from $D_{\mathcal{L},\mathbf{r}/(2g)}$.*

Proof of Proposition 4.11. The proof follows from both parts in a straightforward manner. We start with a radius $r \geq \sqrt{2q}\eta_\epsilon(\mathcal{I})$ and samples from $D_{\mathcal{I},\mathbf{r}}$ for each $\mathbf{r} \in W_{r,\zeta,T}$, and we define $r' = r\omega(1)/(\sqrt{2}\alpha q) > \sqrt{2n}/\lambda_1(\mathcal{I}^\vee)$ for some function $\omega(1)$ that grows asymptotically faster than a constant. The goal is to obtain samples of $D_{\mathcal{I},\mathbf{r}'}$ for some $\mathbf{r}' \geq r'$ of our choice; let $\mathbf{r}' = r' \cdot \mathbf{t}$ where $\mathbf{t} \geq 1$ is a scaled representation of \mathbf{r}' .

Using the samples and the RLWE oracle, Proposition 4.12 yields an algorithm that solves $R\text{-GDP}_{\mathcal{I}^\vee,g}$ for $g = o(1)\alpha q/(\sqrt{2}r)$. Then, Proposition 3.10 yields an algorithm that, on input \mathbf{t} , outputs an independent sample from $D_{\mathcal{I},\mathbf{t}/(2g)}$. Plugging in the values of g and r , we have

$$\frac{\mathbf{t}}{2g} = \frac{r}{o(1)\sqrt{2}\alpha q} \mathbf{t} = \frac{r' \sqrt{2}\alpha q}{\omega(1)o(1)\sqrt{2}\alpha q} \mathbf{t} = \xi r' \mathbf{t} = \mathbf{r}',$$

where $\xi = 1/(\omega(1)o(1))$ vanishes since $\omega(1)$ and $o(1)$ are the inverse of each other. Finally, the condition on the quantum part holds since

$$g = \frac{o(1)\alpha q}{\sqrt{2}r} = \frac{o(1)\omega(1)}{2r'} < \frac{\lambda_1(\mathcal{I}^\vee)}{\sqrt{8n}}.$$

□

The rest of the section is dedicated to the proof of the classical part, that we divide in two components. The first one is the generation of Ring-LWE samples. The procedure takes an instance $y \in R^\vee$ of $R\text{-BDD}_{\mathcal{I}^\vee, d}$ where $y = x + e$ for $\|e\|_\infty \leq d$ and produces Ring-LWE samples where the secret is related to x . The reason why we use a BDD and not a GDP instance is that it is more convenient to work with a strict bound for e . Since samples from D_g have ℓ_∞ norm at most $g \cdot \omega(\sqrt{\log n})$ with all but negligible probability, the sample generation can be applied for $R\text{-GDP}_{\mathcal{I}^\vee, g}$ where $d = g \cdot \omega(\sqrt{\log n})$.

In Section 4.2.3, we show how to carry out the sample generation. In Section 4.2.4, we will use it to prove Proposition 4.12 using the method for finding an oracle's center presented in Section 3.3.

4.2.3 From BDD to Ring-LWE samples

Our goal is to convert a $R\text{-BDD}$ instance on the dual ideal \mathcal{I} into RLWE samples, following the proof sketch in Lemma 3.12 very closely. The result is presented below. Combined with Lemma 3.13, it can be seen as a classical part for the Iterative Step which uses a search RLWE oracle, in a similar fashion as Proposition 3.9.

The reduction takes the $R\text{-BDD}_{\mathcal{I}^\vee, d}$ instance and creates samples of $A_{s, \psi}$ using the invertible bijection $\theta_t : R_q \rightarrow \mathcal{I}_q$ given by Lemma 2.44. We structure the proof in three parts. First, we introduce the algorithm to generate the samples. Second, we prove the correctness of the procedure assuming that the errors behave nicely. Third, we prove that the error distributions do behave as claimed. We closely follow the proofs of Lemmas 4.7 and 4.8 in [LPR10], but we reformulate the statement and generalize it to the use of samples from an elliptical discrete Gaussian $D_{\mathcal{I}, r}$ instead of a spherical one. This generalization will be required later, and is mentioned (but not presented) in [PRSD17]. We also generalize the proof and do not restrict the parameters to those required by Proposition 4.11.

Lemma 4.13 (From $R\text{-BDD}$ to RLWE samples). *Let $y = x + e \in \mathcal{I}^\vee$ such that $x \in \mathcal{I}^\vee$ and $\|e\|_\infty \leq d$ be an instance of $R\text{-BDD}_{\mathcal{I}^\vee, d}$. There is a probabilistic algorithm that, given such instance, an $r \geq \sqrt{2}q\eta(\mathcal{I})$, the factorization of q , and samples from $D_{\mathcal{I}, r}$ for $r \geq r$, outputs samples from $A_{s, \psi}$ where $\psi = D_{r'}$ for $r'_i = \sqrt{(r_i|\sigma_i(e)|/q)^2 + (rd/q)^2}$, and $s = \theta_t(x \bmod q\mathcal{I}^\vee) = tx \in R_q^\vee$.*

Proof. 1. PROCEDURE. Given y and the samples of $D_{\mathcal{I}, r}$, our algorithm proceeds as follows. First, it computes an element $t \in \mathcal{I}$ such that $t\mathcal{I}^{-1}$ and (q) are coprime ideals, as in Lemma 2.44. Since we know the factorization of q over the integers, we can also factor the ideal (q) and do this efficiently. Second, for each sample $z \in \mathcal{I}$ from $D_{\mathcal{I}, r}$, it samples $e' \leftarrow D_{rd/q}$ and calculates

$$a = \theta_t^{-1}(z \bmod q\mathcal{I}), \quad b = (zy)/q + e' \bmod R^\vee.$$

Recall that the bijection θ_t can be inverted efficiently. Finally, it outputs $(a, b) \in R_q \times \mathbb{T}$, which is our Ring-LWE sample.

2. CORRECTNESS. We will now prove that (a, b) is a sample of $A_{s, \psi}$ as claimed. First, we show that $a \in R_q$ follows a distribution which is at negligible distance from uniform. Indeed, we have that $r \geq r \geq \eta_\epsilon(\mathcal{I})$, so Lemma 2.9 implies that the distribution of $z \bmod q\mathcal{I}$ is close to uniform¹ (recall that $z \leftarrow D_{\mathcal{I}, r}$). Since $a = \theta_t^{-1}(z \bmod q\mathcal{I})$ and θ_t is a bijection, then $a \bmod qR$ is also uniform.

¹We need to use the generalization of Lemma 2.9 for elliptical Gaussians, see Definition 2.11. Lemma 2.10 is also generalized in this way.

We can now condition on a fixed value of a and study the distribution of b . Using that $y = x + e$, we have

$$b = (zy)/q + e' = (zx)/q + e(z/q) + e' \bmod R_q^\vee.$$

For the first term zx/q , we have that $z = \theta_t(a) = at \in \mathcal{I}_q$, by definition of a . Since $x \in \mathcal{I}^\vee = \mathcal{I}^{-1}R^\vee$, then $zx = atx \bmod R_q^\vee$ (notice that $zx \in \mathcal{I}_q\mathcal{I}^{-1}R^\vee = R_q^\vee$). Besides, $s = tx \bmod R_q^\vee$, therefore

$$zx/q = as/q \bmod R^\vee.$$

To conclude the proof, we need to prove that the sum of the second and third terms $e(z/q) + e'$ follows the distribution $D_{\mathbf{r}'}$. For the second term $e(z/q)$, we know that z follows a distribution $D_{\mathcal{I},\mathbf{r}}$. Fixed a , z/q follows a distribution $D_{\mathcal{I}+u/q,\mathbf{r}/q}$ where $u = \theta_t(a) \bmod q\mathcal{I}$. Notice that the coset u/q appears because z/q is not a point of the lattice \mathcal{I} . Therefore, we need to analyze the distribution that results from the sum of the elliptical Gaussian distribution $D_{\mathcal{I}+u/q,\mathbf{r}/q}$ and $D_{rd/q}$.

3. ERROR DISTRIBUTIONS. We have that $e \in K$, so we can write $e(z/q) + e'$ as $e(z/q + e'/e)$ and study both terms in the sum. The distribution of e'/e is an elliptical Gaussian $D_{\mathbf{t}}$ where $t_i = rd/q|\sigma_i(e)|$, since the multiplication of embedded elements is coordinate-wise. Besides,

$$t_i = \frac{rd}{q|\sigma_i(e)|} \geq \frac{rd}{q\|e\|_\infty} \geq \frac{r}{q},$$

so we can write $D_{\mathbf{t}}$ as a sum of a continuous Gaussian and an elliptical Gaussian, $D_{\mathbf{t}} = D_{r/q} + D_{\mathbf{t}'}$ where $t_i'^2 = t_i^2 - (r/q)^2 \geq 0$. We now want to study the distribution of $z/q + e'/e$, which corresponds to $D_{\mathcal{I}+u/q,\mathbf{r}/q} + D_{r/q} + D_{\mathbf{t}'}$. For the first two terms, notice that $r/q \geq \sqrt{2}\eta_\epsilon(\mathcal{I})$ by assumption. This means that their widths are above the smoothing condition, so we can apply Lemma 2.10 to obtain that their sum is at negligible distance from the elliptical continuous Gaussian $D_{\sqrt{r^2+r^2}/q}$. We conclude that $z/q + e'/e$ follows a distribution $D_{\sqrt{r^2+r^2}/q} + D_{\mathbf{t}'} = D_{\mathbf{h}}$, which is a continuous elliptical Gaussian where

$$h_i^2 = \frac{r_i^2 + r^2}{q^2} + t_i'^2 = \frac{r_i^2 + r^2}{q^2} + t_i^2 - \frac{r^2}{q^2} = \left(\frac{r_i}{q}\right)^2 + \left(\frac{rd}{q|\sigma_i(e)|}\right)^2.$$

Thus, $e(z/q) + e'$ follows the distribution $\psi = D_{\mathbf{r}'}$ as claimed. \square

As we mentioned, we can interpret this result as the classical part of a reduction to search RLWE. To complete the reduction from R -BDD to SRLWE, notice that we recover the secret $x = \theta_t^{-1}(s \bmod q\mathcal{I}^\vee)$, which is a solution modulo qR^\vee . Then, we can apply Lemma 3.13 to obtain the general solution (i.e., not modular) to the R -BDD problem. The reduction in Lemma 3.13 is lattice-preserving, hence it applies as well to ideal lattices [LPR10].

The equivalence between the BDD and GDP instances was explained before. The key observation is that if we have a R -GDP $_{\mathcal{I}^\vee,g}$ for $g = \alpha q/(\sqrt{2}r)$ and take the bound $d = f(n) \cdot g$ for a function $f(n) = \omega(\sqrt{\log n})$, then our R -GDP $_{\mathcal{I}^\vee,g}$ instances will be R -BDD $_{\mathcal{I}^\vee,d}$ instances with all but negligible probability. Besides, once we apply Lemma 4.13, the error distribution of the samples will be distributed exactly as Γ_α (Definition 4.5). We will need this fact in the coming section.

4.2.4 Using the Decision Oracle

With Lemma 4.13, we finally have all the tools required to prove the classical part of the IS (Proposition 4.12). We want to find the $e \in \mathcal{I}^\vee$ from a R -GDP instance $y = x + e$. For this,

we will construct OHCP instances from oracles that create Ring-LWE samples using Lemma 4.13 and whose hidden center is related to the error vector e . The approach is similar to the proof sketch of Proposition 3.18

The proof that we present follows [PRSD17] very closely, but reformulates some parts (particularly towards the end of the proof) for improved clarity. The algorithm will create $s_1 + s_2 = n - s_2$ oracles \mathcal{O}_i , one for each of the real coordinates of the canonical embedding and one for each conjugate pair of complex embeddings. When testing each of the oracles to approach their center, we leave all the other coordinates fixed. To this respect, the set $W_{r,\zeta,t}$ from Definition 4.10 provides the necessary samples that are used to create the Ring-LWE instances given to the decision oracle.

We remark that the procedure does not aim to find the secret of the samples, which would be equivalent to constructing a reduction from search RLWE to decision RLWE. We simply need the answers of the decision oracle to guide each of our \mathcal{O}_i . Hence, the value of s in the samples does not play an important role and we can randomize it.

Before proving the main result, we introduce a lemma that says that if the product of the coordinates of a vector \mathbf{r} is large, the whole \mathbf{r} must be above the smoothing condition of the lattice. This is a key part of the proof, since it ensures that when we increase a single coordinate of the canonical embedding, our Ring-LWE samples are overridden by the error distribution and indistinguishable from uniform.

Lemma 4.14. *For any $\mathbf{r} \in G$ where $c = (\prod_{i=1}^n r_i)^{1/n} \geq 1$, then $\mathbf{r} \geq \eta_\epsilon(R^\vee)$ for $\epsilon = e^{-c^2 n}$.*

Proof. Consider the lattice \mathcal{L} generated as $T^{-1}\sigma(R^\vee)$ where T is a diagonal matrix such that $T_{i,i} = r_i$. This is, \mathcal{L} is an elliptical scaling of $\sigma(R^\vee)$. We have that $\mathcal{L}^* = T\sigma(R^\vee)^*$ and any $\mathbf{w} \in \mathcal{L}^* \setminus \{0\}$ is of the form $T^{-1}\overline{\sigma(v)}$ for some $v \in R$. Since $|N(v)| = \prod_i |\sigma_i(v)| \geq 1$, then

$$\|\mathbf{w}\|^2 = \sum_{i=1}^n r_i^2 |\sigma_i(v)|^2 \geq n \left(\prod_{i=1}^n r_i^2 |\sigma_i(v)|^2 \right)^{1/n} \geq c^2 n$$

and therefore $\lambda_1(\mathcal{L}^*) \geq c\sqrt{n}$. By Lemma 2.12, this implies that $1 \geq \eta_\epsilon(\mathcal{L})$, which is equivalent to $\mathbf{r} \geq \eta_\epsilon(R^\vee)$. \square

Proof of Proposition 4.12

The proof is structured in four parts, preceded by some useful remarks. First, we present the procedure and the construction of the OHCP instances. Second, we prove that the procedure is correct (assuming the validity of the OHCP instances) and that it returns the error term $e \in R^\vee$, solving the target R -GDP instance. Third, we prove that the OHCP instances are valid according to Definition 3.16. Fourth, we prove that the conditions in Theorem 3.17 are met.

We introduce three preliminary observations. The first is that it suffices to solve R -GDP $_{\mathcal{I}^\vee, g'}$ with non-negligible probability, where $g' = \alpha q / (\sqrt{2}r)$ (i.e., $g' = g/o(1)$), thanks to the self-reducibility result in Proposition 2.24. The second is that we will always work with $\epsilon \leq 2^{-n}$, concerning $\eta_\epsilon(\mathcal{I})$ and the statistical distances from uniform distributions. Hence, we can use the convenient inequality $\eta_\epsilon(\mathcal{I}) \geq O(1)\sqrt{n}/\lambda_1(\mathcal{I}^\vee)$ from Lemma 2.12. The third is that we can assume that $\alpha \geq e^{-n}$. Otherwise we would have, due to the previous fact, that with overwhelming probability

$$\|\sigma(e)\| \leq \sqrt{n}g \leq \alpha\sqrt{n}/\eta_\epsilon(\mathcal{I}) \leq 2^{-n}\lambda_1(\mathcal{I}^\vee)$$

and the problem can be solved using Babai's algorithm (Lemma 2.23). Here, we have used that $g = o(1)\alpha q / \sqrt{2}r \leq \alpha/\eta_\epsilon(\mathcal{I})$ since $r \geq \sqrt{2}q\eta_\epsilon(\mathcal{I})$.

1. **PROCEDURE.** Let $\kappa = \text{poly}(n) \geq 100n^2m$ such that the advantage of the decision $\text{RLWE}_{q,\Gamma_\alpha}$ oracle requiring m samples is at least $2/\kappa$ (we can always choose a κ since the oracle has a non-negligible advantage). The algorithm takes all inputs stated in Proposition 4.12 with the difference that our instance $y = x + e$ is from $R\text{-GDP}_{\mathcal{I}^\vee, g'}$, where $g' = \alpha q/(\sqrt{2}r)$. The desired output is the offset e .

The procedure simulates $s_1 + s_2$ decision oracles $\mathcal{O}_i : \mathbb{R} \times \mathbb{R}^+$ for $i = 1, \dots, s_1$ and $\mathcal{O}_i : \mathbb{C} \times \mathbb{R}^+$ for $i = s_1 + 1, \dots, s_1 + s_2$, following the definitions in Section 3.3, and where $z \in \mathbb{C}$ is embedded in a vector $\mathbf{z} \in \mathbb{R}^2$ as usual. The oracles are constructed as follows. Let $\tau_i : \mathbb{R} \mapsto K_{\mathbb{R}}$ defined as $\sigma_i^{-1}(ze_i)$ for $i = 1, \dots, s_1$ and $\tau_i : \mathbb{C} \mapsto K_{\mathbb{R}}$ as $\sigma_i^{-1}(ze_i + \bar{z}e_{i+s_2})$ for $i = s_1 + 1, \dots, s_1 + s_2$, where e_i is a vector of the canonical basis that is 1 in the i -th coordinate and 0 elsewhere. The oracle $\mathcal{O}_i(z, t)$ takes fresh samples from $D_{\mathcal{I}, \mathbf{r}_{ij}}$ where $\mathbf{r}_{i,j} \in W_{r,\zeta,T} \subset G$ following the notation in Definition 4.10, and such that $(1 + \zeta)^j = e^t$. Then, it creates a $R\text{-GDP}_{\mathcal{I}^\vee, g'}$ instance $y' = y - \tau_i(z)$ (i.e., $y' = x + e - \tau_i(z)$) and it transforms the $D_{\mathcal{I}, \mathbf{r}_{ij}}$ samples and y' into a set $M_{i,z,t}$ of samples from $A_{s,\psi}$. This is done using the algorithm in Lemma 4.13 with a bound $d = f(n)g'$ for a fixed $f(n) = \omega(\sqrt{\log n}) \leq n$. Finally, the oracle \mathcal{O}_i calls the $\text{RLWE}_{q,\Gamma_\alpha}$ oracle on input $M_{i,z,t}$ and returns the same answer.

We will prove that each of the \mathcal{O}_i have a hidden center $\sigma_i(e)$, and that their acceptance probability on input (z, t) depends only on $e^t|z - \sigma_i(e)|$. Besides, they are valid instances of $(e^{-\kappa}, e^{-\kappa}, 1 + 1/\kappa)$ -OHCP with a distance parameter $d' = d/(1 + 1/\kappa)$. Hence, the algorithm runs the procedure in Theorem 3.17 to find a good approximation to the center $z_i \approx \sigma_i(e)$. Finally, it runs Lemma 2.23 on $x' = y - \sum_{i=1}^{s_1+s_2} \tau_i(z_i)$ and outputs $x \in \mathcal{I}^\vee$ (or, equivalently, e).

2. **CORRECTNESS.** We will assume that the OHCP instances are valid and prove that the above algorithm outputs the correct e with all but negligible probability. Let z_i be the approximations of $\sigma_i(e)$. By Theorem 3.17, we have that

$$|z_i - \sigma_i(e)| \leq e^{-\kappa} d' \leq e^{-\kappa} f(n)g' \leq e^{-\kappa} f(n)\alpha/\eta_\epsilon(\mathcal{I}) \leq 2^{-n} \lambda_1(\mathcal{I}^\vee)/\sqrt{n}$$

where we again used that $g' = \alpha q/\sqrt{2}r \leq O(1)\alpha/\eta_\epsilon(\mathcal{I})$, besides the fact that $e^{-\kappa} f(n)\alpha \leq e^{-2n} f(n)\alpha \leq 2^{-n}$. Hence, the error in the approximation is bounded by

$$\left\| e - \sum_{i=1}^{s_1+s_2} \tau_i(z_i) \right\| = \left(\sum_{i=1}^{s_1+s_2} |\sigma_i(e) - z_i|^2 \right)^{1/2} \leq \sqrt{n} \max_i |\sigma_i(e) - z_i| \leq 2^{-n} \lambda_1(\mathcal{I}^\vee).$$

Therefore, Babai's algorithm (Lemma 2.23) will return the correct answer $x \in \mathcal{I}^\vee$.

3. **VALIDITY OF THE OHCP INSTANCES.** We need to prove that the \mathcal{O}_i are valid instances of $(e^{-\kappa}, e^{-\kappa}, 1 + 1/\kappa)$ -OHCP according to Definition 3.16. The distribution of the Ring-LWE samples $M_{i,z,t}$ generated as in Lemma 4.13 only depends on $e^t|z - \sigma_i(e)|$; this is because we are fixing the coordinate i that characterizes the vector $\mathbf{r}' \in W_{r,\zeta,t}$ used to generate the samples, and the non-spherical coordinate i changes by a factor e^t . Hence, p_i depends only on that same factor. Besides, $|z - \sigma_i(e)| \leq d$ and as $d' = f(n)g'$, then $e^{-\kappa} d' \leq |\sigma_i(e)| \leq d'$ with overwhelming probability, since $e^{-\kappa}$ is negligible on n and e is sampled from the Gaussian $D_{g'}$.

4. **CONDITIONS OF THEOREM 3.17.** The *first* condition that we need to prove is that there exists a t^* such that $p_i(0, t^*) - p_\infty \geq 1/\kappa$. Recall that p_∞ is the probability of the oracle accepting on uniform (random) input samples. Notice that $p_i(0, t)$ is identical for each of the \mathcal{O}_i , since the input to the sample generation by Lemma 4.13 is the initial GDP instance y (no offset is added). Besides, if we take $t^* = 0$, then the error distribution of the samples is

Γ_α , so our distinguisher has an advantage greater than $2/\kappa$ by assumption. It follows that $p_i(0,0) - p_\infty \geq 1/\kappa$.

The *second* condition is the convergence of the probability, $|p_i(0,t) - p_\infty| \leq 2e^{-t/\kappa}$ for every t . We will calculate the statistical distance between the samples of $M_{i,0,t}$ (on which \mathcal{O}_i accepts with probability $p_i(0,t)$) and the uniform distribution (respectively p_∞). We know, by Lemma 4.13, that the samples from $M_{i,0,t}$ follow the distribution A_{s,r'_i} for a vector \mathbf{r}'_i given by

$$\begin{cases} r'_{i,i}{}^2 = e^{2t}(r|\sigma_i(e)|/q)^2 + (rd/q)^2 & \geq e^{2t}(r|\sigma_i(e)|/q)^2 \\ r'_{i,j}{}^2 = (r|\sigma_j(e)|/q)^2 + (rd/q)^2 & \geq (r|\sigma_j(e)|/q)^2 \end{cases}.$$

Notice that all coordinates are spherical except for one of them, as in $W_{r,\zeta,t}$. We now choose $\epsilon = e^{-c^2n}$ as in Lemma 4.14 where $c = \prod_{i=1}^n r'_{i,j}$. We know that, since $c \geq 1$ by assumption, then $\mathbf{r}' \geq \eta_\epsilon(R^\vee)$. Thus, by Lemma 2.9, the statistical distance between a sample from $M_{i,0,t}$ (recall that they are produced mod R^\vee) and the uniform distribution is at most $\epsilon/2$. The statistical distance cannot be amplified by any function, so the advantage of the distinguisher cannot be larger. If our decision oracle has access to m samples, we can apply the union bound on m and we have that, for large t such as $t \geq \kappa/10 \geq 10n^2m$,

$$\begin{aligned} |p_i(0,t) - p_\infty| &\leq m\epsilon/2 \\ &\leq (m/2) \exp\left(-n \prod_{j=1}^n (r'_{i,j})^{2/n}\right) \\ &\leq (m/2) \exp\left(-ne^{2t/n}(r/q)^2 \prod_{j=1}^n |\sigma_j(e)|^{2/n}\right). \end{aligned}$$

Since $|\sigma_j(e)| \geq e^{-n}g' \geq e^{-n}\alpha q/(\sqrt{2}r)$ with all but negligible probability and $\alpha \geq 2^{-n}$, then $|\sigma_j(e)| \geq e^{-2n-1/2}(q/r)$. Besides, we can just take that $e^{2t/n-4n-1} \gg t/(\kappa n) + \log(m/4)/n$ because t is very large compared to n . Plugging in both facts yields:

$$\begin{aligned} |p_i(0,t) - p_\infty| &\leq (m/2) \exp\left(-ne^{2t/n}(r/q)^2 [e^{-2n-1/2}(q/r)]^2\right) \\ &\leq (m/2) \exp(-ne^{2t/n-4n-1}) \\ &\leq 2 \exp(-t/\kappa). \end{aligned}$$

For small $t \leq \kappa/10$, we have that $2e^{-t/\kappa} \geq 2e^{-1/10} \geq 1$. As $|p_i(0,t) - p_\infty| \leq 1$, the condition also holds.

The *third* condition follows by studying the statistical distance between the distributions of M_{i,z,t_1} and M_{i,z,t_2} . Notice that these differ only on the i -th coordinate. If $|t_1 - t_2| \leq 1$, then by Lemma 2.7 their statistical distance is bounded by $10m(e^{|t_1-t_2|} - 1) \leq \kappa|t_1 - t_2|$. If $|t_1 - t_2| \geq 1$, then the statistical difference is larger than 1, but the probability difference cannot be larger than $1 \leq \kappa|t_1 - t_2|$. In both cases, $|p_i(t_1) - p_i(t_2)| \leq \kappa|t_1 - t_2|$, so $p_i(z,t)$ is κ -Lipschitz in t . \square

One of the main drawbacks of this result is that it requires a Ring-LWE oracle for the distribution over the error distributions Γ_α , which is a rather restrictive requirement. With a loss of roughly $O((mn)^{1/4})$ in the approximation factor, one can relax this requirement to an oracle for a single spherical Gaussian error distribution D_ξ . The result is the following [PRSD17].

Corollary 4.15 (Main theorem, spherical error). *Let K be a number field of degree n and $R = \mathcal{O}_K$ its ring of integers. Let $q \geq 2$, $\alpha \in (0, 1)$ such that $\alpha q > 2\omega(1)$, $\rho = \omega(\sqrt{\log n}) \cdot (nm/\log(nm))^{1/4}$, and $\xi \geq \rho/q$. There is a polynomial time quantum reduction from $R\text{-DGS}_\gamma$ (for a negligible $\epsilon = \epsilon(n) > 0$) to $\text{RLWE}_{q, D_\epsilon}$, for any*

$$\gamma = \max \left\{ \sqrt{2}\rho\eta_\epsilon(\mathcal{I})/\xi, \sqrt{2n}/\lambda_1(\mathcal{I}^\vee) \right\}.$$

4.3 Related Work and Variants

The Ring-LWE problem is supported by solid hardness results, as we have just seen. Nevertheless, they might be unsatisfactory for some real-world applications. The conjecture that algebraically structured lattices are quantum-safe is at least questionable, and the canonical embedding is less intuitive and implementable than, for instance, the polynomial coefficient embedding. In this final section, we discuss Ring-LWE related work on aspects such as weak instances and variants. The discussion is mostly informal and concludes with a brief survey of open questions.

4.3.1 Weaknesses

The security of Ring-LWE is not as straightforward as the security of LWE, especially when considering applications. As we mentioned previously, the primal RLWE variant (Definition 4.6) is widely used in practice since it is faster and easier to implement. There exist instances of primal RLWE that are vulnerable to number-theoretical attacks. In [ELOS15], samples from polynomial rings $R_q = \mathbb{Z}_q/(f(x))$ are shown to be vulnerable against distinguishing attacks when f has a root that has a small order modulo q . They also provide lower bounds on the width of the error distribution for which primal RLWE is safe. The attack first uses the coefficient embedding (PLWE) and is later extended to a family of number fields for Ring-LWE. In particular, some polynomials of the form $x^n + q - 1$ for a prime q are vulnerable. It is worth noticing that these attacks do not extend to cyclotomic polynomials.

Later, the same family of attacks was made more efficient and extended directly to the search version of primal RLWE in [CIV16]. A general study was done in [Pei16b], concluding that all these weaknesses are related to a poor transformation of the error distribution from the primal to the dual version (recall the discussion that follows Proposition 4.7).

A second threat on Ring-LWE arises in the improvements on algorithms for solving ideal lattice problems, both classical and quantum. When Ring-LWE was first proposed, no quantum algorithm could significantly outperform any classical algorithm on neither lattices or ideal lattices. However, this has changed in the last years, as polynomial-time quantum algorithms for approximating ideal SVP up to subexponential factors have been developed [Duc17]. This creates a hardness gap between unstructured lattice problems (on which the security of LWE is based) and algebraic lattices. Moreover, it questions the assumption that ideal lattice problems are quantum-safe.

In cryptographic applications, subexponential approximations are insufficient to break Ring-LWE, so the loss of security is still not significant. Nevertheless, cryptosystems based on plain LWE such as FrodoKEM [BCD⁺16] are a great backup plan in the event of further progress on ideal lattice algorithms. The security estimates of post-quantum cryptosystems are based on the best available algorithms for solving lattice problems (somehow analogously to how the security of RSA is estimated by the performance of integer factoring algorithms).

4.3.2 Variants of LWE

Besides Ring-LWE, several variants of Learning with Errors have been proposed in recent years². Most of them pursue the goal of being more efficient and easily implementable than the original LWE problem, while maintaining good security properties.

Polynomial-LWE was proposed before the actual Ring-LWE problem in [SSTX09]. The approach is very similar; an instance of Polynomial-LWE is an instance of Ring-LWE in which the elements are considered as polynomials, and they are embedded in a lattice via the coefficient embedding instead of the canonical embedding. This makes error sampling easier because we can directly sample polynomials with small coefficients. The main problematic with the coefficient embedding is that the error distribution is not well controlled for rings other than $\mathbb{Z}[x]/(x^n + 1)$ for n a power of 2. The equivalence of Polynomial-LWE and Ring-LWE was established in [RSW18], where they manage to control the error distribution.

Module-LWE tries to interpolate between plain LWE and Ring-LWE. An sample from a Module-LWE distribution is given by $\mathbf{a} = (a_1, \dots, a_d) \in (R_q)^d$ and $b \in \mathbb{T}$, where the secret $\mathbf{s} \in (R_q)^d$ [LS15]. Instead of having a single ring product $a \cdot s$, now we have two vectors of ring elements. In applications, the vector dimension is usually small; for example Kyber [BDK⁺18] uses $d = 2, 3, 4$ depending on the security level (higher d implies higher security and larger computational cost).

Middle-Product LWE. Let $d \in \mathbb{Z}$, $f, g \in \mathbb{Z}[x]$ two polynomials such that $\deg(f) \leq d_1, \deg(g) \leq d_2$, and assume that $d - 2k = d_1 + d_2 - 1$ for some integer k . The d -middle product between f and g is given by multiplying $f \cdot g$ as usual, deleting all terms of degree up to x^{k-1} , deleting all terms of degree larger than x^{d+k} , and dividing all remaining terms by x^k [RSSS17]. The result is a polynomial of degree d . A sample from a Middle-Product LWE instance consists on replacing the multiplication operation in polynomial-LWE (or in Ring-LWE) by the middle product of two polynomials; the rest remains the same. The efficiency gain is that the middle-product operation can be computed very fast, and that there is no need to work on a quotient ring (i.e., $\mathbb{Z}[x]/(f)$).

Learning with Rounding is a variant of discrete LWE (where $e \in \mathbb{Z}_q$) that replaces the addition of an error $e \leftarrow \chi$ by a deterministic rounding of the product $\langle \mathbf{a}, \mathbf{s} \rangle$. The rounding function divides the $(0, q - 1)$ interval in d parts and rounds the product to the nearest part [BPR12]. Learning with Rounding enjoys a reduction from LWE (for certain parameters) and is used by the NIST PQC finalist SABER [DKRV18]. Some of its main advantages are its simplicity and the reduced amount of randomness required for sample generation.

The web of reductions between some of the variants of Ring-LWE has been greatly simplified recently by Peikert and Pepin, who introduce a *generalized algebraic LWE* problem that generalizes variants such as Module-LWE, Polynomial-LWE or Order-LWE [PP19]³. Then, they achieve tight reductions from Ring-LWE to the generalized problem. Separately, they reduce Ring-LWE for any ring to Middle-Product LWE, managing to narrow the error distributions. In general, their results imply that most of these convenient forms of LWE can benefit from the hardness results for Ring-LWE.

²A schematic view of reductions and equivalences, mostly between primal/dual Ring-LWE and Polynomial-LWE, can be found in [Ros20]. In [PP19], a diagram of their reductions is also presented.

³The revisited version of this paper includes stronger results.

4.3.3 Future Directions

At the time of writing, many questions around LWE and Ring-LWE remain open. Arguably, the main question addressing LWE directly is whether the reductions that we have seen can be made entirely classical, continuing the work in [BLP⁺13]. Such a result would increase the confidence that we have on LWE-based constructions. In parallel, there is ongoing research in finding good (possibly quantum) algorithms for lattice problems. There is limited progress in the search of efficient approximation algorithms for SVP and related problems. Nevertheless, this is not the case for ideal lattices, and the gap that exists between lattices and ideal lattices certainly questions the assumption that ideal lattice problems are hard, and hence the hardness of Ring-LWE [Duc17].

Another line of research concerns the algebraic LWE variants [PP19]. The introduction of more efficient and provably secure variants of LWE and Ring-LWE is of interest since it will improve the quality of our cryptographic constructions. It is also important to close the web of reductions among them. In order to construct practical and secure (such as IND-CCA) cryptosystems on these variants, new tools such as the modular Fujisaki-Okamoto Transform [HHK17] have been introduced.

Regarding practical constructions, proving the security of several variants of NTRU [HPS98], some of which are candidates for standardization by NIST, also remains an open problem [AASA⁺20]. For many other cryptosystems, the security proofs are not completely satisfactory, since the chosen LWE or Ring-LWE parameter sets do not correspond to the parameters required in the reductions of Chapters 3 and 4. Besides, some of them (such as Kyber [BDK⁺18]) introduce “rounding” steps to improve efficiency that are not considered in the security proofs, and others rely on the random oracle model. There is also room for improvement in advanced LWE-based constructions such as fully-homomorphic encryption [Gen09] and attribute-based encryption schemes.

We conclude that the cryptographic community has solid reasons to believe that we can construct schemes based on the Learning with Errors Problem that are against quantum adversaries. Nevertheless, the security of constructions based on Ring-LWE-based is less certain and requires further study. We believe that the intense research activity around LWE will lead to important cryptographic breakthroughs in the coming years.

Bibliography

- [AASA⁺20] Gorjan Alagic, Jacob Alperin-Sheriff, Daniel Apon, David Cooper, Quynh Dang, John Kelsey, Yi-Kai Liu, Carl Miller, Dustin Moody, Rene Peralta, et al. Status report on the second round of the nist post-quantum cryptography standardization process. *US Department of Commerce, National Institute of Standards and Technology*, 2020.
- [Bab86] László Babai. On lovász’lattice reduction and the nearest lattice point problem. *Combinatorica*, 6(1):1–13, 1986.
- [Ban93] Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(1):625–635, 1993.
- [BCD⁺16] Joppe Bos, Craig Costello, Léo Ducas, Ilya Mironov, Michael Naehrig, Valeria Nikolaenko, Ananth Raghunathan, and Douglas Stebila. Frodo: Take off the ring! practical, quantum-secure key exchange from LWE. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pages 1006–1018, 2016.
- [BDK⁺18] Joppe Bos, Léo Ducas, Eike Kiltz, Tancreède Lepoint, Vadim Lyubashevsky, John M Schanck, Peter Schwabe, Gregor Seiler, and Damien Stehlé. Crystalskyber: a cca-secure module-lattice-based kem. In *2018 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 353–367. IEEE, 2018.
- [BLP⁺13] Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 575–584, 2013.
- [BPR12] Abhishek Banerjee, Chris Peikert, and Alon Rosen. Pseudorandom functions and lattices. In David Pointcheval and Thomas Johansson, editors, *Advances in Cryptology – EUROCRYPT 2012*, pages 719–737, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.
- [Cas59] J. W. S. Cassels. *An Introduction to the Geometry of Numbers*. Springer Berlin Heidelberg, 1959.
- [CIV16] Wouter Castryck, Ilia Iliashenko, and Frederik Vercauteren. Provably weak instances of ring-lwe revisited. In Marc Fischlin and Jean-Sébastien Coron, editors, *Advances in Cryptology – EUROCRYPT 2016*, pages 147–167, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg.
- [Con] Kieth Conrad. The different ideal. Available online at: <http://www.math.uconn.edu/~kconrad/blurbs/gradnumthy/different.pdf>.

- [DD12] Léo Ducas and Alain Durmus. Ring-lwe in polynomial rings. In Marc Fischlin, Johannes Buchmann, and Mark Manulis, editors, *Public Key Cryptography – PKC 2012*, pages 34–51, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.
- [DKRV18] Jan-Pieter D’Anvers, Angshuman Karmakar, Sujoy Sinha Roy, and Frederik Vercauteren. Saber: Module-LWR based key exchange, CPA-secure encryption and CCA-secure KEM. In *International Conference on Cryptology in Africa*, pages 282–305. Springer, 2018.
- [DKS98] I. Dinur, G. Kindler, and S. Safra. Approximating-cvp to within almost-polynomial factors is np-hard. pages 99–109. IEEE Comp Soc, 1998.
- [Duc17] Léo Ducas. Advances on quantum cryptanalysis of ideal lattices. *Nieuw Archief voor Wiskunde*, 5:184–189, 2017.
- [ELOS15] Yara Elias, Kristin E Lauter, Ekin Ozman, and Katherine E Stange. Provably weak instances of ring-lwe. In *Advances in Cryptology – CRYPTO 2015*, pages 63–92. Springer, 2015.
- [Gen09] Craig Gentry. Fully homomorphic encryption using ideal lattices. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 169–178, 2009.
- [HHK17] Dennis Hofheinz, Kathrin Hövelmanns, and Eike Kiltz. A modular analysis of the fujisaki-okamoto transformation. In *Theory of Cryptography Conference*, pages 341–371. Springer, 2017.
- [HPS98] Jeffrey Hoffstein, Jill Pipher, and Joseph H Silverman. Ntru: A ring-based public key cryptosystem. In *International Algorithmic Number Theory Symposium*, pages 267–288. Springer, 1998.
- [Kho05] Subhash Khot. Hardness of approximating the shortest vector problem in lattices. *J. ACM*, 52(5):789–808, September 2005.
- [LJ08] Hendrik W Lenstra Jr. Lattices. 2008.
- [LLL82] Arjen K Lenstra, Hendrik Willem Lenstra, and László Lovász. Factoring polynomials with rational coefficients. *Mathematische annalen*, 261:515–534, 1982.
- [LPR10] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In Henri Gilbert, editor, *Advances in Cryptology – EUROCRYPT 2010*, pages 1–23, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.
- [LPR13] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. A toolkit for ring-lwe cryptography. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 35–54. Springer, 2013.
- [LS15] Adeline Langlois and Damien Stehlé. Worst-case to average-case reductions for module lattices. *Designs, Codes and Cryptography*, 75(3):565–599, 2015.
- [MG02] Daniele Micciancio and Shafi Goldwasser. *Complexity of Lattice Problems*. Springer US, 2002.
- [Mil08] James S Milne. *Algebraic number theory*. JS Milne, 2008.

- [MR07] Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on gaussian measures. *SIAM Journal on Computing*, 37(1):267–302, 2007.
- [Pei10] Chris Peikert. An efficient and parallel Gaussian sampler for lattices. In *Annual Cryptology Conference*, pages 80–97. Springer, 2010.
- [Pei16a] Chris Peikert. A decade of lattice cryptography. *Foundations and Trends® in Theoretical Computer Science*, 10(4):283–424, 2016.
- [Pei16b] Chris Peikert. How (not) to instantiate ring-lwe. In Vassilis Zikas and Roberto De Prisco, editors, *Security and Cryptography for Networks*, pages 411–430, Cham, 2016. Springer International Publishing.
- [PP19] Chris Peikert and Zachary Pepin. Algebraically structured lwe, revisited. In *Theory of Cryptography Conference*, pages 1–23. Springer, 2019.
- [PR07] Chris Peikert and Alon Rosen. Lattices that admit logarithmic worst-case to average-case connection factors. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 478–487, 2007.
- [PRSD17] Chris Peikert, Oded Regev, and Noah Stephens-Davidowitz. Pseudorandomness of ring-LWE for any ring and modulus. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 461–473, 2017.
- [Reg09] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *Journal of the ACM (JACM)*, 56(6):1–40, 2009.
- [Reg10] Oded Regev. The learning with errors problem. *Invited survey in CCC*, 7:30, 2010.
- [Rib13] Paulo Ribenboim. *Classical theory of algebraic numbers*. Springer Science & Business Media, 2013.
- [Ros20] Miruna Rosca. *On algebraic variants of Learning With Errors*. PhD thesis, Université de Lyon, 2020.
- [RSSS17] Miruna Roșca, Amin Sakzad, Damien Stehlé, and Ron Steinfeld. Middle-product learning with errors. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology – CRYPTO 2017*, pages 283–297, Cham, 2017. Springer International Publishing.
- [RSW18] Miruna Rosca, Damien Stehlé, and Alexandre Wallet. On the ring-lwe and polynomial-lwe problems. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 146–173. Springer, 2018.
- [SSTX09] Damien Stehlé, Ron Steinfeld, Keisuke Tanaka, and Keita Xagawa. Efficient public key encryption based on ideal lattices. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 617–635. Springer, 2009.
- [Ste14] Damien Stehlé. Advanced cryptographic primitives course 3: The learning with errors problem, 2014. <http://perso.ens-lyon.fr/damien.stehle/downloads/Course3.pdf>.