A provably secure variant of NTRU cryptosystem

Danilo Ciaffi

Advised by **Guilhem Castagnos**
Alessandro Languasco

Università di Padova

Université de Bordeaux

Academic year 2017-2018

Cheesy catchphrase

Contents

Introduction

As the development of quantum computing starts to seem closer, it appears imperative to find new protocols able to stay sound under a potential quantum attack. In fact, most traditional methods heavily rely on factorization or discrete logarithm, and a polynomial-time quantum algorithm is known for both. On the other hand, some problems arising from lattices seem to be difficult both from a classical and quantum point of view.

In addition to the conjectured quantum resistance, lattice-based schemes yield some other interesting properties:

- they are simple to implement and highly parallelizable: due to the very nature of lattices, the operations involved are usually only sums or matrix-vector multiplication. On top of that, these operations are modulo a relatively small integer, giving an even stronger bound to running times;
- they usually enjoy strong security guarantees from worst-case hardness. This means that breaking their security is proved at least as hard as solving some lattice problems in any of its instances, including the worst ones;

making them appear as very desirable and viable alternatives to traditional methods.

In 1996 Hoffstein, Pipher ad Silverman presented NTRUEncrypt [HPS98], which is to date the fastest known lattice-based encryption scheme. Its moderate key-sizes, excellent asymptotic performance and conjectured resistance to quantum attacks make it a perfect candidate to succeed where factorization and discrete log fail. Unfortunately, no security proof has been produced for NTRUEncrypt nor for its signature counterpart NTRUSign.

In 2013 Stehlé and Steinfield in [SS11] proposed to apply some mild modification to the encryption and signature scheme to make them provably secure in the standard (resp. random oracle) model, under the assumed quantum (resp. classical) hardness of standard worst-case lattice problems, restricted to a family of lattices related to some cyclotomic fields. In particular they showed that if the secret key polynomials of the encryption scheme are chosen from discrete Gaussians, then the public key, *i.e* their ratio, is statistically indistinguishable from uniform. The security will then follow from the hardness of the R-LWE problem, proved in [LPR12] and described in Chapter 2.

The aim of this thesis is to present [SS11] in a slightly more accessible form, providing some more background and details in some points. On the other hand, a basic knowledge of algebraic number theory is taken for granted, and sometimes, to make the work more digestible to the reader, some not-strictly-necessary or rather technical proofs and details have been flew over in Chapter 2.

The outline of this work is the following:

- Chapter 1 will be devoted to all the necessary preliminaries;
- Chapter 2 to the presentation of R-LWE problem;
- Chapter 3 to the actual main results in Stehlé's and Stenfield's paper.

Notation

Before getting started, let's fix some notation:

- If q is a non-zero integer, and $(R, +, \times)$ a ring, we let R_q denote the R/qR and $R^\times$ the set of invertible elements of R.
- If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\|\mathbf{x}\|$ will always denote the Euclidean norm of $\mathbf{x}$ and $\langle \mathbf{x}, \mathbf{y} \rangle$ the Euclidean inner product of $\mathbf{x}$ and $\mathbf{y}$.
- If E is a set, we let $U(E)$ denote the uniform distribution over E.
- We will write $z \sim D$ when the random variable z is sampled from the distribution D.
- Given two functions $f, g : \mathbb{N} \to \mathbb{R}$ we will use the following notations:
  - $f(n) = O(g(n))$ if there exist some $k > 0$ and $n_0 > 0$ such that for any $n \geq n_0$, $|f(n)| \leq k \cdot |g(n)|$;
  - $f(n) = \Theta(g(n))$ if there exist $k_1, k_2 > 0$ and $n_0 > 0$ such that for any $n \geq n_0$, $k_1 \cdot g(n) \leq f(n) \leq k_2 \cdot g(n)$;
  - $f(n) = \omega(g(n))$ if there exist $k > 0$ and $n_0 > 0$ such that, for any $n \geq n_0$, $|f(n)| \geq k \cdot |g(n)|$;
  - $f(n) = \Omega(g(n))$ if there exist $k > 0$ and $n_0 > 0$ such that, for all $n \geq n_0$, $f(n) \geq k \cdot g(n)$.

  We will also say $f(n) = \widetilde{O}(g(n))$ (or $\widetilde{\Theta}(g(n)), \widetilde{\omega}(g(n)), \widetilde{\Omega}(g(n))$) if $f(n) = O(g(n))$ (or $\Theta(g(n)), \omega(g(n)), \Omega(g(n))$ respectively) up to a $\log(n)$ factor.
- A function $f(n)$ is said *negligible* if $f(n) = n^{-\omega(1)}$ and a sequence of events E_n holds with *overwhelming* probability if $\Pr[\neg E_n] \leq f(n)$ for a negligible function f.

  In practice, we consider negligible an amount $< 2^{-30}$.
- We will say a cryptosystem has n bits of security when on average at least 2^n operations are required to break it.
- K will be used for number fields, $\mathcal{O}_K$ for the ring of integers of K.

Chapter 1

Preliminaries

In this chapter we collect some results that are crucial to the understanding of what follows. We will start by giving the definitions and briefly illustrating some of the properties of lattices, to move to the depiction of the problems that make them interesting for cryptographic purposes. After that some probability and number theory tools will be given, since they are necessary to describe LWE problem. These can be found in literature in [Reg09], [HPSS08], [Pei16]. We will continue by illustrating the NTRU cryptosystem in a slightly different form than the original, as presented in [MR08], and conclude by providing some technical results on random q-ary lattices, where q is an integer number.

1.1 Lattices

1.1.1 Definitions and first properties

Definition 1.1.1. An n-dimensional (full-rank) *lattice* L is the free abelian group generated by a basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ of $\mathbb{R}^n$. The integer n is called the *dimension* of the lattice.

The set $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is still called a *basis* of L and can be written in the form of a matrix $B = [\mathbf{b}_1, \dots, \mathbf{b}_n] \in \mathbb{R}^{n \times n}$ whose columns are the vectors of the basis. From this we can obtain the lattice generated by B as $L(B) = \{B\mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^n\}$. The *fundamental domain* of L is $\mathcal{F}(L) = \{\sum_{i=1}^n t_i \mathbf{b}_i \mid t_i \in [0, 1]\}$ and its volume is a constant of the lattice, called $\det(L)$.

Remark 1.1.2. Once a basis is given, another can be obtained through an invertible matrix with integer coefficients (*i.e.* an element of $GL_n(\mathbb{Z})$), which has determinant ±1. Though very rigid, these transformations are of great interest, as the problems we are going to see in the next section can be very hard or very easy depending on the used basis. Typically we will call a “good” basis one composed by short and almost orthogonal vectors (according to the euclidean norm).

Definition 1.1.3. The *minimum* of a lattice L is the euclidean norm of any of its non-zero shortest vectors, namely the positive real number

$$\lambda_1(L) := \min\{\|\mathbf{x}\| \mid \mathbf{x} \in L\}.$$

This notion can be generalized to define the *i-th successive minimum* $\lambda_i(L)$ as the smallest $r \in \mathbb{R}$ such that L has exactly i linearly independent vectors of length at most r.

Definition 1.1.4. The *dual lattice* $L^\vee$ of a lattice L is

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

Definition 1.1.5. A lattice $L \subseteq \mathbb{Z}^n$ is said *q-ary* for some integer q if $q\mathbb{Z} \subseteq L$.

Let us denote $R = \mathbb{Z}[x]/\Phi$, where $\Phi \in \mathbb{Z}[x]$ is a monic irreducible polynomial of degree n. For $\mathbf{a} \in R_q^m$, consider the following families of R-modules:

$$\begin{aligned} L(\mathbf{a}) &= \{(t_1, \dots, t_m) \in R^m \mid t_i = a_i s \bmod q \text{ for } i = 1, \dots, m \text{ and } s \in R_q\}\,; \\ \mathbf{a}^\perp &= \left\{(t_1, \dots, t_m) \in R^m \mid \sum a_i t_i = 0 \pmod q\right\}. \end{aligned}$$

These correspond to mn-dimensional lattices via the map sending an element of R^m to the concatenation of the vectors of coefficients. Since these lattices are obviously q-ary, they are called *module q-ary lattices.*

1.1.2 Computational problems

Most of the time, proving the security of a cryptosystem means to show that breaking it is as hard as solving some computational problem known - or assumed - to be hard. Here we present those problems arising from lattices that are useful to our purposes.

Definition 1.1.6. Given an arbitrary basis B of a lattice $L = L(B)$, the *Shortest Vector Problem* (SVP) consists in finding a shortest non-zero lattice vector, *i.e.* a vector $\mathbf{v} \in L$ such that $\|\mathbf{v}\| = \lambda_1(L)$.

Definition 1.1.7. Given an arbitrary basis B of a lattice $L = L(B)$ and a point $\mathbf{x}$ in $\mathbb{R}^n$, the *Closest Vector Problem* (CVP) consists in finding the lattice vector whose distance from $\mathbf{x}$ is minimal.

In most practical applications, we use an approximation of this problems for the average case to worst case reductions. In particular, these instances are parametrized by an approximation factor $\gamma \geq 1$, usually polynomial in the dimension n of the lattice.

Definition 1.1.8. Given an arbitrary basis B of an n-dimensional lattice $L = L(B)$, the *Approximate Shortest Vector Problem* (SVP_γ) consists in finding a non-zero lattice vector $\mathbf{v}$ such that $\|\mathbf{v}\| \leq \gamma(n) \cdot \lambda_1(L)$.

Definition 1.1.9. Given a basis B of an n-dimensional lattice $L = L(B)$, the *Approximate Shortest Independent Vector Problem* (SIVP_γ) requires to find a set $S = \{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$ of n linearly independent lattice vectors with $\|\mathbf{s}_i\| \leq \gamma(n) \cdot \lambda_n(L)$ for all $i = 1, \ldots, n$.

A key point in the next chapter will be the possibility to reduce from the search problems defined in Definition 1.1.8 and 1.1.9 to the relative decision problem.

Definition 1.1.10. Given an arbitrary basis B of an n-dimensional lattice $L = L(B)$, the *Decisional Approximate SVP* (GapSVP_γ) consists in distinguishing whether $\lambda_1(L) \leq 1$ or $\lambda_1(L) > \gamma(n)$.

Remark 1.1.11. We can see that GapSVP_γ is a promise problem, *i.e* it is a decision problem in which the “yes” and “no” instances do not exhaust the set of all possible inputs. In particular, nothing can be said if $1 < \lambda_1(L) \leq \gamma$, therefore the choice of the parameter becomes crucial.

The last problem we present is very important for the Learning with Errors problem and asks to find the unique lattice vector that is the closest to a given point $\mathbf{t} \in \mathbb{R}^n$, which is known to be “sufficiently” close to the lattice.

Definition 1.1.12. Given a basis B of an n-dimensional lattice $L = L(B)$, and a target point $\mathbf{t} \in \mathbb{R}^n$ such that $dist(\mathbf{t}, L) < d = \lambda_1(L)/(2\gamma(n))$, the *Bounded Decoding Distance* problem (BDD_γ) consists in finding the unique lattice vector $\mathbf{v} \in L$ such that $\|\mathbf{t} - \mathbf{v}\| < d$.

To the best of our knowledge, when the approximation factor γ is polynomial in n the problems presented so far turn out to be intractable both with a classical and a quantum approach. Therefore, it is conjectured that there is no polynomial-time classical or quantum algorithm that solves worst-case approximated lattice problems when $\gamma = poly(n)$.

1.2 Discrete Gaussian distributions

Many modern cryptographic protocols make use of a discrete form of the Gaussian distribution over lattices, called *discrete Gaussian distribution*. Here we present the results we are going to need in the rest of the treatise.

Definition 1.2.1. For any positive integer n, vector $\mathbf{c} \in \mathbb{R}^n$ and real $s > 0$, the *Gaussian function* $\rho_{s,\mathbf{c}} : \mathbb{R}^n \to \mathbb{R}^+$ of parameter (or width) s and centered in $\mathbf{c}$ is defined as

$$\rho_{s,\mathbf{c}}(\mathbf{x}) := \exp\left(\frac{-\pi\|\mathbf{x} - \mathbf{c}\|^2}{s^2}\right),$$

where $\|\mathbf{x}\|$ denotes the euclidean norm of $\mathbf{x}$.

Since the total measure associated to ρ_r is $\mathbb{R}$ $\int_{\mathbf{x}\in\mathbb{R}^n} \rho_{s,\mathbf{c}}(\mathbf{x})\mathrm{d}\mathbf{x} = s^n$, by normalizing we get the *continuous Gaussian distribution*, defined by the probability density,

$$D_{s,\mathbf{c}}(\mathbf{x}) := \frac{\rho_{s,\mathbf{c}}(\mathbf{x})}{s^n}.$$

We will omit s if $s = 1$ and $\mathbf{c}$ if $\mathbf{c} = \mathbf{0}$.

Remark 1.2.2. It's easy to see that ρ_s is invariant under rotations of $\mathbb{R}^n$ and that $\rho_s(\mathbf{x}) = \prod_{i=1}^n \rho_s(x_i)$. This means that a sample from the Gaussian distribution D_s can be obtained by taking n independent samples from the 1-dimensional Gaussian distribution.

Definition 1.2.3. For any countable set A and any real parameter (width) $s > 0$, the *discrete Gaussian probability distribution* $D_{A,s}$ is defined as

$$\forall \mathbf{x} \in A,\ D_{A,s}(\mathbf{x}) := \frac{\rho_s(\mathbf{x})}{\rho_s(A)};$$

with $\rho_s(A) = \sum_{\mathbf{x}\in A} \rho_s(\mathbf{x})$.

As we will be working with number fields, let s_1 and s_2 be natural numbers, we introduce

$$H = \{(x_1, \ldots, x_n) \in \mathbb{R}^{s_1} \times \mathbb{C}^{2s_2} \mid x_{s_1+s_2+j} = \overline{x_{s_1+j}}\ \forall j \leq s_2\} \subseteq \mathbb{C}^n.$$

Let $n = s_1 + 2s_2$ and let us define the following basis $\{\mathbf{h}_i\}_{i\leq n}$ of H:

$$\begin{cases} \mathbf{h}_j = \mathbf{e}_j & \text{for } j \leq s_1 \\ \mathbf{h}_j = \frac{1}{\sqrt{2}}(\mathbf{e}_j + \mathbf{e}_{j+s_2}) & \text{for } s_1 < j \leq s_1 + s_2 \\ \mathbf{h}_j = \frac{i}{\sqrt{2}}(\mathbf{e}_{j-s_2} - \mathbf{e}_j) & \text{for } s_1 + s_2 < j \leq s_1 + 2s_2 = n. \end{cases}$$

This basis makes H isomorphic to $\mathbb{R}^n$ as an inner product space and allows us to give the following definition:

Definition 1.2.4. Given $\mathbf{r} = (r_1, \ldots, r_n) \in (\mathbb{R}^+)^n$, with $n = s_1 + 2s_2$ and such that $r_{j+s_1+s_2} = r_{j+s_1}$ for each $j \in \{1, \ldots, s_2\}$, a sample from the *elliptical Gaussian distribution* $D_{\mathbf{r}}$ is given by $\sum_{i=1}^n x_i \mathbf{h}_i$, where each x_i is chosen independently from the 1-dimensional Gaussian distribution D_{r_i} over $\mathbb{R}$.

Definition 1.2.5. For a lattice L and a real $\epsilon > 0$, the *smoothing parameter* $\eta(L)$ is the smallest $\lambda > 0$ such that $\rho_{1/\lambda}(L^\vee \setminus \{\mathbf{0}\}) \leq \epsilon$.

Definition 1.2.6. Let D_1, D_2 be two probability density functions on $\mathbb{R}^n$; the *statistical distance* between them is

$$\Delta(D_1, D_2) := \int_{\mathbb{R}^n} |D_1(\mathbf{x}) - D_2(\mathbf{x})| \mathrm{d}\mathbf{x}$$

in the continuous case and

$$\Delta(D_1, D_2) := \sum_{\mathbf{x} \in S} |D_1(\mathbf{x}) - D_2(\mathbf{x})|$$

in the case of a discrete set S.

Here we state some technical lemmas that are more or less classical in literature, see *e.g.* [Ban93],[HPSS08], [Pei16].

Lemma 1.2.7. *Let $f : \mathbb{R}^n \to \mathbb{C}$ be a function and $\widehat{f}$ denote its Fourier transform. Then for any lattice L, it holds that $f(L) = det(L^\vee)\widehat{f}(L^\vee)$.*

Lemma 1.2.8. *For any lattice L, positive real $s > 0$ and vector $\mathbf{c}$, we have $\rho_{s,\mathbf{c}}(L) \leq \rho_s(L)$.*

Lemma 1.2.9. *For any full-rank lattice $L \subseteq \mathbb{R}^n$ and $\delta > 0$, we have*

$$\eta_\delta(L) \leq \sqrt{\frac{\log(2n(1+1/\delta))}{\pi}} \lambda_n(L).$$

Proof. Let $s = \sqrt{\frac{\log(2n(1+1/\delta))}{\pi}}$. We want to show that $\rho_{1/s}(L^\vee \setminus \{\mathbf{0}\}) \leq \delta$.

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be linearly independent vectors in $\mathbb{R}^n$ of length at most $\lambda_n(L)$ and define the set $S_{i,j} = \{\mathbf{x} \in L^\vee \mid \langle \mathbf{v}_i, \mathbf{x} \rangle = j \in \mathbb{Z}\}$. For any fixed i these sets form a partition of $L^\vee$, and since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent, every $\mathbf{x} \in L^\vee$ must have non-zero product with at least one of the vectors. Hence $L^\vee \setminus \{\mathbf{0}\} = \bigcup_{i=1}^n (L^\vee \setminus S_{i,0})$.

For every $i = 1, \ldots, n$, let $\mathbf{u}_i = \mathbf{v}_i / \|\mathbf{v}_i\|^2$ be a vector of length $1/\|\mathbf{v}_i\| \geq 1/\lambda_n(L)$ pointing the same direction as $\mathbf{v}_i$. For all $j \in \mathbb{Z}$, we obtain

$$\rho_{1/s}(S_{i,j}) = \exp(-\pi \|js\mathbf{u}_i\|)^2 \rho_{1/s}(S_{i,0} - j\mathbf{u}_i)$$

and since $S_{i,j} - j\mathbf{u}_i$ is a shift of the set $S_{i,0}$, there exists some vector $\mathbf{w}$ (orthogonal to $\mathbf{u}_i$) such that $S_{i,j} - j\mathbf{u}_i = S_{i,0} - \mathbf{w}$. Now, by the previous lemma

$$\rho_{1/s}(S_{i,j} - j\mathbf{u}_i) = \rho_{1/s}(S_{i,0} - \mathbf{w}) = \rho_{1/s,\mathbf{w}}(S_{i,0}) \leq \rho_{1/s}(S_{i,0})$$

and, using $\|\mathbf{u}_i\| \geq 1/\lambda_n(L)$ and the bound $\sum_{j \neq 0} x^{-j^2} \leq 2\sum_{j>0} x^{-j} = 2/(x-1)$ (for $x > 1$), we get

$$\begin{aligned}
\rho_{1/s}(L^\vee \setminus S_{i,0}) &= \sum_{j \neq 0} \rho_{1/s}(S_{i,j}) \\
&\leq \sum_{j \neq 0} e^{-\pi(s/\lambda_n)^2 j^2} \rho_{1/s}(S_{i,0}) \\
&\leq \frac{2}{e^{\pi(s/\lambda_n)^2} - 1} \rho_{1/s}(S_{i,0}) \\
&= \frac{2}{e^{\pi(s/\lambda_n)^2} - 1} (\rho_{1/s}(L^\vee) - \rho_{1/s}(L^\vee \setminus S_{i,0})).
\end{aligned}$$

Since $\rho_{1/s}$ is positive, we can write

$$\rho_{1/s}(L^\vee \setminus \{\mathbf{0}\}) \leq \sum_{j\neq 0} \rho_{1/s}(L^\vee \setminus S_{i,0}) \leq \frac{2n}{e^{\pi(s/\lambda_n)^2}+1}\rho_{1/s}(L^\vee)$$

and, using $\rho_{1/s}(L^\vee) = 1 + \rho_{1/s}(L^\vee \setminus \{\mathbf{0}\})$, we get

$$\rho_{1/s}(L^\vee \setminus \{\mathbf{0}\}) \leq \frac{2n}{e^{\pi(s/\lambda_n)^2}+1-2n} < \frac{2n}{e^{\pi(s/\lambda_n)^2}-2n} := \delta,$$

which concludes the proof. □

Lemma 1.2.10. *Let B denote the unitary ball centered in 0. For each $t \geq (2\pi)^{-1/2}$ and $\mathbf{u} \in \mathbb{R}^n$, one has:*

1. $\rho(L \setminus t\sqrt{n}B) < (t\sqrt{2\pi e}\, e^{-\pi t^2})^n \rho(L)$,
2. $\rho(L+\mathbf{u}) \setminus \sqrt{n}B < 2(t\sqrt{2\pi e}\, e^{-\pi t^2})^n \rho(L)$.

Lemma 1.2.11. *For any n-dimensional full-rank lattice L, $\mathbf{c} \in \mathbb{R}^n$ and reals $\sigma \geq \eta_\delta(L)$, $\delta \in (0,1)$, we have*

$$\Pr_{\mathbf{x}\sim D_{L,\sigma,\mathbf{c}}}\left[\|\mathbf{x}-\mathbf{c}\| > \sigma\sqrt{n}\right] \leq \frac{1+\delta}{1-\delta}\cdot 2^{-n}.$$

Proof. It is enough to prove the statement for $\sigma = 1$. Let us write

$$\Pr_{\mathbf{x}\sim D_{L,\sigma,\mathbf{c}}}\left[\|\mathbf{x}-\mathbf{c}\| > \sigma\sqrt{n}\right] = \frac{\rho((L-\mathbf{c}) \setminus \sqrt{n}B)}{\rho_{\mathbf{c}}(L)},$$

where B indicates the unitary ball centered in the origin. By Lemma 1.2.10, with a constant $t = 1$ the numerator is bounded by $2^{-n}\rho(L)$, and by Lemma 1.2.7 we get

$$\begin{aligned}\rho_{\mathbf{c}}(L) &= \det(L^\vee)\widehat{\rho_{\mathbf{c}}}(L^\vee)\\ &= \det(L^\vee)\sum_{\mathbf{y}\in L^\vee}\widehat{\rho_{\mathbf{c}}}(\mathbf{y})\\ &= \det(L^\vee)\sum_{\mathbf{y}\in L^\vee} e^{-2\pi i\langle \mathbf{c},\mathbf{y}\rangle}\widehat{\rho}(\mathbf{y})\\ &= \det(L^\vee)(1+\varepsilon)\end{aligned}$$

where $|\varepsilon| \leq |\rho(L^\vee) \setminus \{\mathbf{0}\}| \leq \delta$. Hence we have $\rho_{\mathbf{c}}(L^\vee) \geq \det(L^\vee)(1-\delta)$, $\rho(L) \leq \det(L^\vee)(1+\delta)$ and thus $2^{-n}\rho(L)/\rho_{\mathbf{c}}(L) \leq 2^{-n}\frac{1+\delta}{1-\delta}$. □

Corollary 1.2.12. *For any full-rank lattice $L \subseteq \mathbb{R}^n$, $\mathbf{c} \in \mathbb{R}^n$, $\delta \in (0,1)$ and $\sigma \geq \eta_\delta(L)$, we have $\rho_{\sigma,\mathbf{c}'}(L') = \frac{\sigma^n}{\det(L)}(1+\varepsilon)$ for some $|\varepsilon| < \delta$, and therefore*

$$\frac{\rho_{\sigma,\mathbf{c}}(L)}{\rho_\sigma(L)} \in \left[\frac{1-\delta}{1+\delta}, 1\right].$$

Corollary 1.2.13. *For any n-dimensional full-rank lattice L, $\mathbf{c} \in \mathbb{R}^n$, reals $\delta \in (0,1)$, $\sigma \geq 2\eta_\delta(L)$ and $\mathbf{b} \in L$ we have*

$$D_{L,\sigma,\mathbf{c}}(\mathbf{b}) \leq \frac{1+\delta}{1-\delta} \cdot 2^{-n}.$$

Corollary 1.2.14. *Let $L' \subseteq L \subseteq \mathbb{R}^n$ be full-rank lattices. Then for any $\mathbf{c} \in \mathbb{R}^n$, $\delta \in (0,1/2)$ and $\sigma \geq \eta(L')$,*

$$\Delta(D_{L,\sigma,\mathbf{c}} \mod L', U(L/L')) \leq 2\delta,$$

where $D_{L,\sigma,\mathbf{c}} \mod L'$ means that the samples of $D_{L,\sigma,\mathbf{c}}$ (which are elements of L) are then reduced modulo L'.

Lemma 1.2.15. *There exists a polynomial-time algorithm that takes as input any basis $\{\mathbf{b}_i\}$ of any lattice $L \subseteq \mathbb{Z}^n$ and $\sigma = \omega(\sqrt{\log n}) \max \|\mathbf{b}_i\|$ and returns samples from a distribution whose statistical distance to $D_{L,\sigma}$ is negligible with respect to n.*

Lemma 1.2.16. *For any full-rank lattice $L \subseteq \mathbb{R}^n$, $\mathbf{c} \in \mathbb{R}^n$, $\delta \in (0,1)$ and $\sigma \geq \eta_\delta(L)$, $t \geq \sqrt{2\pi}$, $\sigma \geq t/\sqrt{2\pi}$ and unit vector $\mathbf{u} \in \mathbb{R}^n$, we have*

$$\Pr_{\mathbf{b} \sim D_{L,\sigma,\mathbf{c}}} \left[|\langle \mathbf{b} - \mathbf{c}, \mathbf{u} \rangle| \leq \frac{\sigma}{t} \right] \leq \frac{1+\delta}{1-\delta} \frac{\sqrt{2\pi e}}{t}.$$

Moreover, if $\sigma \geq \eta_\delta(L)$,

$$\Pr_{\mathbf{b} \sim D_{L,\sigma,\mathbf{c}}} \left[|\langle \mathbf{b} - \mathbf{c}, \mathbf{u} \rangle| \geq \sigma t \right] \leq \frac{1+\delta}{1-\delta} t e^{-\pi t^2} \sqrt{2\pi e}.$$

Proof. Let U be an orthonormal matrix, $\mathbf{u}^T$ its first row and choose $\mathbf{b}' \sim D_{L',\sigma,\mathbf{c}'}$, with $L' = UL$ and $\mathbf{c}' = U\mathbf{c}$. Defining X as the random variable that corresponds to the first component of $\mathbf{b}' - \mathbf{c}'$ we have

$$\Pr \left[|X| \leq \frac{\sigma}{t} \right] = \frac{(\rho_{\sigma,\mathbf{c}'} \cdot \mathbb{1}_{\sigma/t,\mathbf{c}'})(L')}{\rho_{\sigma,\mathbf{c}'}}$$

where $\mathbb{1}_{\sigma/t,\mathbf{c}'}(\mathbf{x})$ has value 1 if $|x_1 - c_1'| \leq \sigma/t$ and 0 otherwise. To estimate the denominator we use that $\eta_\delta(L') = \eta_\delta(L)$ and $\det(L') = \det(L)$, so by Corollary 1.2.12 we have $\rho_{\sigma,\mathbf{c}'}(L') = \frac{\sigma^n}{\det(L)}(1+\varepsilon)$ for some $|\varepsilon| < \delta$. For the numerator, for any $\mathbf{x} \in \mathbb{R}^n$ we have $\mathbb{1}_{\sigma/t,\mathbf{c}'}(\mathbf{x}) \leq e^{K(1 - \frac{|x_1 - c_1'|^2}{\sigma^2/t^2})}$ for a constant $K = \frac{1}{2} - \frac{\pi}{t^2} \in [0, \frac{1}{2}]$. As a consequence we have

$$(\rho_{\sigma,\mathbf{c}'} \cdot \mathbb{1}_{\sigma/t,\mathbf{c}'})(L') \leq e^K \rho_{\sigma,D\mathbf{c}'}(DL'),$$

where D is a matrix with the upper left element equals to $\sqrt{1 + \frac{Kt^2}{\pi}}$ and is the identity elsewhere. A straight-forward computation shows that $\det(DL') =$

$\sqrt{1+\frac{Kt^2}{\pi}}\det(L') = \sqrt{1+\frac{Kt^2}{\pi}}\det(L)$ and $\eta_\delta(DL') \leq \sqrt{1+\frac{Kt^2}{\pi}}\eta_\delta(L')$, therefore σ respects the hypotheses and we can apply Corollary 1.2.12 again to get

$$\begin{aligned}
\Pr\left[|X| \leq \frac{\sigma}{t}\right] &= \frac{(\rho_{\sigma,\mathbf{c}'} \cdot \mathbb{1}_{\sigma/t,\mathbf{c}'})(L')}{\rho_{\sigma,\mathbf{c}'}} \\
&\leq \frac{e^K \rho_{\sigma,D\mathbf{c}'}(DL')}{\frac{\sigma^n(1+\varepsilon)}{\det(L)}} \\
&\leq \frac{e^{\frac{1}{2}}e^{-\frac{\pi}{t^2}}}{\sqrt{1+\frac{Kt^2}{\pi}}\det(L')}\frac{1+\varepsilon}{1-\varepsilon} \\
&\leq \frac{e^{\frac{1}{2}}e^{-\frac{\pi}{t^2}}}{\sqrt{1+\frac{Kt^2}{\pi}}\det(L')}\frac{1+\delta}{1-\delta}
\end{aligned}$$

Now by the hypotheses on t we have that $e^{-\frac{\pi}{t^2}} < 1$, and since $\sqrt{1+\frac{Kt^2}{\pi}} = \sqrt{1+\left(\frac{1}{2}-\frac{\pi}{t^2}\right)\frac{t^2}{\pi}} = t/\sqrt{2\pi}$ we obtain

$$\frac{e^{\frac{1}{2}}e^{-\frac{\pi}{t^2}}}{\sqrt{1+\frac{Kt^2}{\pi}}\det(L')}\frac{1+\delta}{1-\delta} \leq \frac{1+\delta}{1-\delta}\frac{\sqrt{2\pi e}}{t},$$

which is the result we wanted.

The proof of the second statement is completely analogous to the first one. □

1.3 Ideal lattices

In this section we present a special class of lattices, that correspond to ideals in the ring of integers of a number field.

Definition 1.3.1. Let K be a number field of degree n, $R = \mathcal{O}_K$ and let σ be any additive injective map $\sigma : R \to \mathbb{C}^n$. The family of ideal lattices for the ring R and the embedding σ is the set of lattices $\sigma(I)$ for integral ideals $I \subseteq R$.

Traditionally, when we work with lattices this embedding is the componentwise immersion, sending the polynomial $f = \sum f_i x^i$ to the component vector $f = (f_1, \ldots, f_n)$, which, by the definition of number field, is an element of $\mathbb{R}^n$. In order to present the Ring-Learning With Errors problem, though, we are going to consider the canonical embedding, more often used in algebraic number theory.

Given any number field $K = \mathbb{Q}(\zeta)$ of degree n we can consider n field homomorphisms $\sigma_i : K \to \mathbb{C}$ that fix any element of $\mathbb{Q}$ and map ζ to each of its conjugates. We indicate the number of real embeddings with s_1 and the number of pairs of complex embeddings with s_2, hence $n = s_1 + 2s_2$. Moreover we can sort them in such a way that σ_j with $j \leq s_1$ is a real embedding and for $j > s_1$ the embeddings are the complex ones and $\sigma_{j+s_1+s_2} = \overline{\sigma_{j+s_1}}$.

Definition 1.3.2. The *canonical embedding* is the map

$$\begin{aligned} \sigma : K &\longrightarrow \mathbb{R}^{s_1} \times \mathbb{C}^{2s_2} \\ x &\longmapsto \sigma(x) = (\sigma_1(x), \ldots, \sigma_n(x)). \end{aligned}$$

This canonical embedding has the property that both addition and multiplication are component-wise. Moreover, due to the pairing of the complex embeddings, σ maps to the space H defined in the previous section and given an integral ideal I with $\mathbb{Z}$-basis $\{u_1, \ldots, u_n\}$, we have that $\sigma(I) \subseteq H$ is an ideal lattice with basis $\{\sigma(u_1), \ldots, \sigma(u_n)\}$.

Another advantage using this embedding is that it allows to think of the Elliptical Gaussian Distribution as a distribution over $K_{\mathbb{R}} := K \otimes_{\mathbb{Q}} \mathbb{R}$, identifying $K_{\mathbb{R}}$ with H and defining the distribution $D_{\mathbf{r}}$ of $x \otimes s \in K_{\mathbb{R}}$ as the distribution $D_{\mathbf{r}}$ of $\sigma(x)s \in H$ with $r'_i = r_i|\sigma_i(x)|$.

1.4 NTRU cryptosystem

NTRU is a public key cryptosystem introduced in [HPS98] by Hoffstein, Pipher and Silverman, and it is up-to-date the most efficient lattice-based cryptosystem and the most used in practice. Even though no proof of security supporting NTRU is known, the inefficiency of the best currently known attacks seems to suggest confidence in the security of the scheme. We will first present it in its original form, but then we will also show a different interpretation provided in [MR08], *i.e.* we will see it as a particular instance of the more general GGH framework.

1.4.1 The original NTRU

Let us fix a prime number n, two integers p and q such that $\gcd(N, q) = \gcd(p, q) = 1$ and denote $R = \mathbb{Z}[x]/(x^n - 1)$.

Definition 1.4.1. For any positive integers d_1 and d_2, we define

$$\mathcal{T}(d_1, d_2) = \left\{ f \in R : \begin{array}{l} f \text{ has } d_1 \text{ coefficients equal to } 1, \\ f \text{ has } d_2 \text{ coefficients equal to } -1, \\ f \text{ has all other coefficients equal to } 0 \end{array} \right\}$$

as the set of *ternary polynomials.*

The NTRUEncrypt key with (public) parameters (N, p, q, d) is built as follows: we first randomly choose two polynomials

$$f \in \mathcal{T}(d+1, d) \quad \text{and} \quad g \in \mathcal{T}(d, d),$$

where f is discarded and resampled until it is invertible both in R_p and R_q.

We store f as the secret key sk and then compute f_p[1] and f_q, the inverses of f respectively in R_p and R_q, to obtain the public key pk:

$$h = f_q g \in R_q.$$

Remark 1.4.2. A polynomial $f \in \mathcal{T}(d, d)$ is never invertible in R_q, because for such f we have $f(1) = 0$, so $\gcd(f, x^n - 1) = x - 1 \neq 1$.

Algorithm 1 Encryption Key Generation

```
Input: $N$, $p$ primes, $q, d$ positive integers, with $\gcd(p, q) = \gcd(N, q) = 1$.
Choose $f \in \mathcal{T}(d+1, d)$ invertible both in $R_p$ and $R_q$
Choose $g \in \mathcal{T}(d, d)$
Compute $f_q$, the inverse of $f$ in $R_q$
Compute $f_p$, the inverse of $f$ in $R_p$
Return secret key $sk = f$ and public key $pk = h = f_q g$
Output: The key pair $(sk, pk)$
```

Let us now encode our plaintext through a polynomial $m \in R$ whose coefficients all lie in the interval $[-p/2, p/2]$, *i.e.* the center-lift of an element in R_p. To encrypt our message we randomly choose $s \in \mathcal{T}(d, d)$ and compute

$$c = phs + m \mod q.$$

Proposition 1.4.3 (NTRU Decryption). *If the NTRUEncrypt parameters* (N, p, q, d) *satisfy*

$$q > (6d + 1)p,$$

then it is always possible to recover the message m from the ciphertext c.

Proof. Let the ciphertext be $c = phs + m \mod q$ as above. To decrypt the message we first multiply both sides by f:

$$fc = pgs + fm \mod q.$$

We now want to reduce further modulo p, but we first have to make sure $fc \mod q$ is the same as fc in R. To prove this, we need to bound its coefficients when computed before the reduction modulo q. Let us use the notation $\|f\|_\infty := \max_{0 \leq i \leq n}\{f_i\}$.

[1] f_p is not really needed for the key generation, but, since it will be used in the decryption process, it is usually stored at this stage to gain in efficiency.

By construction we have $g, s \in \mathcal{T}(d, d)$, so $\|gs\|_\infty \leq 2d$. On the other hand, $f \in \mathcal{T}(d+1, d)$ and the coefficients of m are in $[-p/2, p/2]$, so $\|fm\|_\infty \leq (2d+1)\frac{p}{2}$. Combining the two we obtain

$$\|pgs + fm\|_\infty \leq p\|gs\|_\infty + \|fm\|_\infty \leq p \cdot 2d + (2d+1)\frac{p}{2} = p\left(3d + \frac{1}{2}\right).$$

Our assumption implies that all the coefficients of fc are also smaller than $q/2$, which means we can look at it as an element in R rather than R_q. We can now finally reduce modulo p:

$$(fc \mod q) \mod p = fm \mod p,$$

and multiplying by f_p we get $m \mod p$. Now being all the coefficients of m in $[-p/2, p/2]$, we have that $m \mod p$ is exactly m. □

Remark 1.4.4. The bound in the result above is actually very strong, because it is very unlikely to have the coefficients line up in such a way to reach the maximum in the products. For this reason, much smaller values of q are used in practice, chosen to verify that the decryption failure probability is smaller than 2^{-80}.

Algorithm 2 Encryption and Decryption

```
Encryption
Input: The NTRU parameters $(N, p, q, d)$, the public key $pk = h$, a message $m \in R$ with coefficients in $[-p/2, p/2]$
  Choose random $s \in \mathcal{T}(d, d)$
  Compute $c = phs + M \mod q$
Output: The ciphertext $c$

Decryption
Input: The NTRU parameters $(N, p, q, d)$, the secret key $sk = f$ (and eventually $f_p$), the ciphertext $c$.
  Compute $fc = pgs + fm \mod q$
  Reduce both sides modulo $p$: $(fc \mod q) \mod p = fm \mod p$
  Compute $f_p fm \mod p$
Output: The plaintext $m$
```

1.4.2 The GGH/HNF public key cryptosystem

The GGH cryptosystem was proposed by Goldreich, Goldwasser, and Halevi in [GGH97], and it is the analogue for lattices of the McEliece cryptosystem in [McE78] which was based on the hardness of decoding linear codes over finite fields. We will here present it quickly:

- The private key is a "good" lattice basis B. Good basis have the property to make the solution of CVP easy, provided that the target s very close to the lattice.

- The public key is a "bad" lattice basis H of the same lattice $L(B) = L(H)$. In a sense, the "worst" possible basis is the Hermite Normal Form (HNF) of B, as it can be efficiently computed from any basis B' of $L(B)$. Moreover any attack on the HNF public key can be easily adapted to work with B' simply by first computing H from B' itself.

- The encryption consists in encoding the message into a short noise vector $\mathbf{r}$ and adding it to chosen lattice point $\mathbf{v}$.

- The decryption requires to find the lattice point $\mathbf{v}$ which is closest to the ciphertext $\mathbf{c} = (\mathbf{r} \bmod H) = \mathbf{v} + \mathbf{r}$, to get the error vector $\mathbf{r} = \mathbf{c} - \mathbf{v}$.

- The signature is obtained applying Babai's round-off CVP approximation algorithm[2] to get a lattice vector close to the message $\mathbf{m}$: $\mathbf{s} = B\lfloor B^{-1}\mathbf{m}\rceil$. To verify the signature, one needs to check that $\mathbf{s} \in L(H)$ and compute the distance $\|\mathbf{s} - \mathbf{m}\|$ to assure that it is sufficiently small.

The correctness of this cryptosystem relies on the fact that the error vector $\mathbf{r}$ is short enough for the lattice point $\mathbf{v}$ to be recovered from $\mathbf{c}$ using the private basis B, *e.g.*, by using Babai's round-off method, $\mathbf{v} = B\lfloor B^{-1}(\mathbf{v} + \mathbf{r})\rceil$.

On the other hand, the security depends on the assumption that solving this instance of the closest vector problem in $L(B) = L(H)$ is computationally hard without knowing of B.

Even though no asymptotically good attack to GGH is known, some attacks break the cryptosystem in practice for moderately large values of the security parameter. This can be avoided by making the security parameter even bigger, but that makes the cryptosystem impractical. In fact, $\Omega(n^2)$ storage is needed for the lattice basis, so the encryption/decryption running times also grow quadratically in the security parameter. This issue raises the need of a more compact representation, and it will be addressed in the next section.

1.4.3 NTRU cryptosystem as a GGH method

Let T be the linear transformation that given a vector $\mathbf{v} = (v_1, \dots, v_n)$ cyclically permutes its entries, *i.e.* $T\mathbf{v} = (v_2, \dots, v_n, v_1)$. For every vector

[2] Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be a basis for a full rank lattice in $\mathbb{R}^n$ and a target $\mathbf{w} \in \mathbb{R}^n$. We can write $\mathbf{w} = \sum_{i=1}^n l_i \mathbf{b}_i$ with $l_i \in \mathbb{R}$ where the l_i's are found computing the vector $\mathbf{l} = (l_1, \dots, l_n) = B^{-1}\mathbf{w}$. Babai's round-off consists in setting $\mathbf{v} = \sum_{i=1}^n \lfloor l_i \rceil \mathbf{b}_i$, (where $\lfloor x \rceil$ indicates the closest integer to x) *i.e.* computing $\mathbf{v} = B\lfloor B^{-1}\mathbf{w}\rceil$. This procedure can be performed using any basis for the lattice, but it works only for a "good" basis.

$\mathbf{v} \in \mathbb{Z}^n$ define $T^*\mathbf{v} = [\mathbf{v}, T\mathbf{v}, \dots, T^{n-1}\mathbf{v}]$ to be the circulant matrix of $\mathbf{v}$. NTRU uses $2n$-dimensional lattices satisfying the following properties:

- they are closed under the linear transformation that maps the vector $(\mathbf{x}, \mathbf{y})$ (where $\mathbf{x}$ and $\mathbf{y}$ are n-dimensional vectors) to $(T\mathbf{x}, T\mathbf{y})$, *i.e.*, the vector obtained by cyclically rotating the coordinates of $\mathbf{x}$ and $\mathbf{y}$ individually;
- they are q-ary lattices, so the membership of $(\mathbf{x}, \mathbf{y})$ in the lattice only depends on $(\mathbf{x}, \mathbf{y})$ mod q.

The parameters of the system are a prime number n and three integers q, p and d_f, and it works as follows:

- *private key*: the private key is a short vector $(\mathbf{f}, \mathbf{g}) \in \mathbb{Z}^{2n}$. To this vector one associates a lattice $\Lambda_q = L((T^*\mathbf{f}, T^*\mathbf{g})^T)$, which is the smallest lattice with the properties above to contain $(\mathbf{f}, \mathbf{g})$. For a correct and efficient functioning of the public key computation, encryption and decryption, these vectors must comply to the following restriction:
  - the matrix $[T^*\mathbf{f}]$ shall be invertible modulo q;
  - the secret vectors $\mathbf{f} \in \mathbf{e}_1 + \{p, 0, -p\}^n$ and $\mathbf{g} \in \{p, 0, -p\}^n$ are randomly chosen in such a way that $\mathbf{f} - \mathbf{e}_1$ and $\mathbf{g}$ have exactly $d_f + 1$ positive entries, d_f negative ones and all others will be zeros.
- *public key*: in accordance with the general GGH/HNF case, the public key for NTRU corresponds to the HNF of the lattice Λ_q defined by the private key. Due to the properties of such lattice and the restriction we imposed on the choice of $\mathbf{f}$, the public key ends up looking as the following:

$$\begin{bmatrix} I & 0 \\ \mathbf{h} & qI \end{bmatrix},$$

  where $\mathbf{h} = [T^*\mathbf{f}]^{-1}\mathbf{g}$. Therefore it can be represented in a compact way through the only vector $\mathbf{h} \in \mathbb{Z}^{2n}$.

- *encryption*: first the message is encoded as a vector $\mathbf{m} \in \{1, 0, -1\}^n$ with exactly $d_f + 1$ positive entries and d_f negative ones. It is then concatenated to a randomly chosen vector $\mathbf{r} \in \{1, 0, -1\}^n$ also with exactly $d_f + 1$ positive entries and d_f negative ones, so in the end we have a short vector $(-\mathbf{r}, \mathbf{m}) \in \{1, 0, -1\}^{2n}$. We can consider this as the error vector used in the general GGH case and reduce it modulo the public basis H to get a new vector $(\mathbf{0}, (\mathbf{m} + [T^*\mathbf{h}]\mathbf{r}) \pmod{q})$. Since the first n coordinates of such vector are always 0, we can store the ciphertext using only the n remaining entries and get $\mathbf{c} = \mathbf{m} + [T^*\mathbf{h}]\mathbf{r} \pmod q$.

- *decryption*: for the decryption, we first notice that for any vectors $\mathbf{f}, \mathbf{h}$ holds $[T^*\mathbf{f}][T^*\mathbf{h}] = [T^*([T^*\mathbf{f}]\mathbf{h})]$. Thus we perform

  $$[T^*\mathbf{f}]\mathbf{c} \bmod q = [T^*\mathbf{f}]\mathbf{m} + [T^*\mathbf{f}][T^*\mathbf{h}]\mathbf{r} \bmod q = [T^*\mathbf{f}]\mathbf{m} + [T^*\mathbf{g}]\mathbf{r} \bmod q$$

  and since all the coordinates of $[T^*\mathbf{f}]\mathbf{m} + [T^*\mathbf{g}]\mathbf{r}$ are bounded in absolute value by $q/2$, so they are also the exact integers values. To conclude, we reduce

  $$[T^*\mathbf{f}]\mathbf{m} + [T^*\mathbf{g}]\mathbf{r} \mod p = I \cdot \mathbf{m} + \mathbb{0} \cdot \mathbf{r} = \mathbf{m}.$$

  It is important to remark that we need $d_f < (q/2 - 1)/(4p) - (1/2)$ for the bound on the coordinates to hold, even though the decryption might work with high probability for larger values of d_f.

Like for GGH, no security proof is known for NTRU, and the confidence in the scheme is due to the inefficiency of the currently known attacks.

Remark 1.4.5. In the table, we find listed some of the current set of suggested parameters for NTRU. The security is expressed in bits, where k bits of security means that the best known attack needs to perform at least the equivalent of 2^k NTRU encryptions operation. On the other hand, the parameter d_f is chosen in such a way that a honest user decryption errors happen with probability less than 2^{-k}.

Estimated security (bits)	n	q	d_f	key size (bits)
80	257	2^{10}	77	2570
80	449	2^8	24	3592
256	797	2^{10}	84	7970
256	14303	2^8	26	114424

We are not going to dive any deeper into the standard choice of parameters, but more on the the topic and the table above can be found in [HG07].

1.5 Additional results

Here we introduce some more technicalities that will be necessary in Chapter 3.

1.5.1 Random q-ary lattices

In this section we are going to generalize the defiintions of $\mathbf{a}^\perp$ and $L(\mathbf{a})$ to comprehend also the ideals of $R_q = \mathbb{Z}[x]/\langle q, \Phi\rangle$. Let $\Phi = \prod_{i<k_q} \Phi_i$ be the factorization in irreducible factors modulo q, where in fact all the factors have the same degree $d_q = n/k_q$. Any ideal of R_q is of the form

$$I_S := \left(\prod_{i\in S} \Phi_i\right) R_q = \{a \in R_q \mid \forall i \in S, a = 0 \bmod \Phi_i\}, \text{with } S \subseteq \{1, \ldots, k_q\}$$

and we can call L_S the lattice corresponding to the ideal $\langle q, \prod_{i\in S} \Phi_i\rangle$, *i.e.* $L_S = \{x \in R \mid (x \mod q) \in I_S\}$.

Given $\mathbf{a} \in R_q^m$ let us define the following families of R-modules:

$$\mathbf{a}^\perp(I_S) := \Big\{(t_1, \ldots, t_m) \in R^m \mid \forall i, (t_i \bmod q) \in I_S, \sum_i t_i a_i = 0 \bmod q\Big\},$$

$$L(\mathbf{a}^\perp, I_S) := \big\{(t_1, \ldots, t_m) \in R^m \mid \exists s \in R_q, \forall i, (t_i \mod q) = a_i \cdot s \bmod I_S\big\},$$

where $S \subseteq \{1, \ldots, k_q\}$. We remark that $\mathbf{a}^\perp(I_S)$ is the intersection of $\mathbf{a}^\perp$ with the cartesian product of m copies of L_S, and that if $S = \emptyset$ (resp. $S = \{1, \ldots, n\}$) then $\mathbf{a}^\perp(I_S) = \mathbf{a}^\perp$ (resp. $L(\mathbf{a}^\perp, I_S) = L(\mathbf{a})$).

We are now going to show that these two modules are one the dual of the other. In the ring R we have $x^{-1} = -x^{n-1}$, so the map $R \to R, a(x) \mapsto a^\star(x) = a(x^{-1})$ is a ring automorphism. This map induces a bijection from the set of factors Φ_i to itself and it has a useful matrix interpretation: let A denote the $n \times n$ matrix having as its i-th row the coefficient vector of $x^i \cdot a(x)$ for $i = 0, \ldots, n-1$, then $a^\star(x)$ has coefficient vector the first column of A. For an ideal $I_S = \left(\prod_{i\in S} \Phi_i\right) R_q$ of R, we let $I_S^\star$ denote the ideal $\left(\prod_{i\in S} \Phi_i^\star\right) R_q$.

Lemma 1.5.1. *Let $S \subseteq \{1, \ldots, k_q\}$, $\overline{S}$ be the complement of S, $\mathbf{a} \in R_q^m$ and $\mathbf{a}^\star \in R_q^m$ be defined by $a_i^\star = a_i(x^{-1})$, for all $i \leq m$. Then considering both sets mn-dimensional lattices:*

$$\mathbf{a}^\perp(I_S)^\vee = \frac{1}{q} L(\mathbf{a}^\star, I_{\overline{S}}^\star).$$

Proof. Let us first prove that $\frac{1}{q}L(\mathbf{a}^\star, I_{\overline{S}}^\star) \subseteq \mathbf{a}^\perp(I_S)^\vee$. Take $\mathbf{t} = (t_1, \ldots, t_m) \in \mathbf{a}^\perp(I_S)$ and $\mathbf{u} = (u_1, \ldots, u_m) \in L(\mathbf{a}^\star, I_{\overline{S}}^\star)$, write $t_i = \sum_{j<n} t_{i,j} x^j$ and $u_i = \sum_{j<n} u_{i,j} x^j$ for any $i \leq m$, we want to show that $\sum_{i\leq m, j\leq n} t_{i,j} u_{i,j} = 0 \bmod q$. This is equivalent to show that the constant coefficient of the polynomial $\sum_{i\leq m} t_i u_i^\star$ is 0 modulo q, therefore it is enough to show that $\langle \mathbf{t}, \mathbf{u}^\star\rangle = 0 \bmod q$.

By definition of the u_i's, there exists $s \in R_q$ such that $(u_i \mod q) = a_i^\star s + b_i$ for some $b_i \in I_{\overline{S}}^\star$. Hence, modulo q,

$$\langle \mathbf{t}, \mathbf{u}^\star\rangle = s^\star \langle \mathbf{t}, \mathbf{a}\rangle + \langle t, \mathbf{b}^\star\rangle = 0,$$

where $\mathbf{b} = (b_1, \ldots, b_m)$. The previous equality holds because $\langle \mathbf{t}, \mathbf{a} \rangle = 0 \bmod q$ by the definition of $\mathbf{t}$ and $\langle \mathbf{t}, \mathbf{b}^\star \rangle = 0 \bmod q$ because $(t_i \bmod q) \in I_S$ and $b_i^\star \in I_{\overline{S}}$ for each $i \leq m$. Thanks to this, we have the inclusion we wanted.

The inverse inclusion $\frac{1}{q} L(\mathbf{a}^\star, I_{\overline{S}}^\star) \supseteq \mathbf{a}^\perp(I_S)^\vee$ is equivalent by duality to $\frac{1}{q} L(\mathbf{a}^\star, I_{\overline{S}}^\star)^\vee \subseteq \mathbf{a}^\perp(I_S)$, and to show the latter, we just have to consider the elements of $L(\mathbf{a}^\star, I_S)$ corresponding to $s = 1$ and repeat a process analogous to the previous inclusion. □

The next step is to show that for a uniformly chosen $\mathbf{a} \in (R_q^\times)^m$, the lattice $L(\mathbf{a}, I_S)$ is extremely unlikely to contain unusually short vectors for the infinity norm, *i.e.* remarkably shorter than the Minkowski upper bound $\det(L(\mathbf{a}, I_S))^{\frac{1}{mn}} = q^{(1-\frac{1}{m})\frac{|S|}{k_q}}$ on $\lambda_1^\infty(L(\mathbf{a}, I_S))$. Observe that we have that $\det(L(\mathbf{a}, I_S)) = q^{(m-1)|S|d_q}$ because there are $q^{|S|d_q + m(n-|S|d_q)}$ points of $L(\mathbf{a}, I_S)$ in the cube $[0, q-1]^{mn}$.

We are then going to give two lower bounds for short vectors: the first lower bound is useful for all parameter settings and matches the Minkowski upper bound up to a factor $\frac{1}{\sqrt{n}} q^{-\varepsilon}$ for an arbitrarily small constant $\varepsilon > 0$; the second is specific to the case $|S| = k_q$ and matches the Minkowski bound up to a factor $q^{-k_q}\varepsilon$, improving on the first by a factor $\approx \sqrt{n}$ in the case $k_q = O(1)$.

Lemma 1.5.2. *Let $n \geq 8$ be a power of 2 and $q \geq 5$. Assume that $\Phi = x^n + 1$ splits into k_q distinct irreducible factors modulo q of degree $d_q = n/k_q$. Then, for $m \geq 2$ and $\varepsilon > 0$, we have*

$$\lambda_1^\infty(L(\mathbf{a}, I_S)) \leq \begin{cases} \frac{1}{\sqrt{n}} q^{(m-1)\frac{|S|}{d_q} - \varepsilon} & \textit{for any } 0 \leq |S| \leq k_q \\ q^{1-\frac{1}{m} - k_q \varepsilon} & \textit{for } |S| = k_q \end{cases}$$

with probability greater than $1 - 2^{4mn} q^{-\varepsilon mn}$ over the uniformly random choice of $\mathbf{a} \in (R_q^\times)^m$.

Proof. By the Chinese Remainder Theorem, we know that R_q and $R_q^\times$ are isomorphic respectively to $(\mathbb{F}_{q^{d_q}})^{k_q}$ and $(\mathbb{F}_{q^{d_q}}^\times)^{k_q}$ through the isomorphism $t \mapsto (t \mod \Phi_i)_{i \leq k_q}$. Let $\Phi_S = \prod_{i \in S} \Phi_i$: it is a generator of I_S of degree $|S|d_q$.

Let p denote the probability over the randomness of $\mathbf{a}$ that $L(\mathbf{a}, I_S)$ contains a non-zero vector $\mathbf{t}$ whose infinity norm is strictly smaller than B. We can give an upper bound for p using the union bound, summing the probabilities $p(\mathbf{t}, s) = \Pr_{\mathbf{a}}[\forall i, t_i = a_i s \mod I_S]$ over all possible values for $\mathbf{t}$ of infinity norm less than B and $s \in R_q/I_S$. Since the a_i's are independent, we have $p(\mathbf{t}, s) = \prod_{i \leq m} p_i(t_i, s)$, where $p_i(t_i, s) = \Pr_{a_i}[t_i = a_i s \mod I_S]$.

If $\gcd(s, \Phi_S) \neq \gcd(t_i, \Phi_S)$, there must be some $j \leq n$ such that either $t_i \bmod \Phi_j = 0$ and $s \mod \Phi_j \neq 0$, or $t_i \mod \Phi_j \neq 0$ and $s \mod \Phi_j = 0$. In both cases, we have $p_i(t_i, s) = 0$ because $a_i \in R_q^\times$, therefore we can assume, without losing generality, that $\gcd(s, \Phi_S) = \gcd(t_i, \Phi_S)$ (up to multiplication by an element of $\mathbb{F}_{q^{d_q}}^{\times}$).

We now assume that $\gcd(s, \Phi_S) = \gcd(t_i, \Phi_S) = \Phi_{S'}$ for some $S' \subseteq S$ of cardinality $0 \leq k \leq |S|$. For any $j \in S'$, we have $t_i = a_i s = 0 \mod \Phi_j$ regardless of the value of $a_i \mod \Phi_j$, while for $j \in S \setminus S'$, we have $s \neq 0 \mod \phi_j$ and there exists a unique value of $a_i \mod \Phi_j$ such that $t_i = a_i s \mod \Phi_j$. Moreover, for any $j \notin S$, the value of $a_i \mod \Phi_j$ can be arbitrary in $\mathbb{F}_{q^{d_q}}^{\times}$. So in the end there are $(q^{d_q} - 1)^{k_q + k - |S|}$ distinct a_i's in $R_q^{\times}$ such that $t_i = a_i s \mod I_S$. This means that $p_i(t_i, s) = (q^{d_q} - 1)^{k - |S|}$. Therefore we can give the following bound for the probability p:

$$p \leq \sum_{0 \leq k \leq |S|} \sum_{\substack{S' \subseteq S \\ |S'| = k}} \sum_{\substack{s \in R/I_S \\ \Phi_{S'} | s}} \sum_{\substack{\mathbf{t} \in R_q^m \\ \forall i, 0 < \|ti\|_\infty < B \\ \forall i, \Phi_{S'} | t_i}} (q^{d_q} - 1)^{m(k - |S|)}.$$

For $|S'| = k$, let $N(B, k)$ denote the number of $\mathbf{t} \in R_q$ such that $\|t\|_\infty < B$ and $\mathbf{t} = \Phi_{S't'}$ for some $\mathbf{t}' \in R_q$ of degree less than $n - kd_q = n(1 - k/k_q)$. We are now going to consider two upper bounds for $N(B, k)$, from which we obtain the claimed bounds on $\lambda_1^\infty(L(\mathbf{a}, I_S))$.

First, for $B = \frac{1}{\sqrt{n}} q^\beta$, we say that $N(B, k) \leq 2^{2n} q^{(\beta - k/k_q)n}$ for $k < \beta k_q$ and 0 otherwise. For this, we observe that $N(B, k)$ is the number of points of the lattice $I_{S'} + q\mathbb{Z}^n = \langle \Phi_{S'}, q \rangle$ in the hypercube $C(2B) = \{\mathbf{v} \in \mathbb{R}^n \mid \|\mathbf{v}\|_\infty < B\}$. Let's denote $\lambda := \lambda_1^\infty(I_{S'} + q\mathbb{Z}^n)$. If we center a hypercube $C(\lambda)$ on each of the $N(B, k)$ points of $I_{S'} + q\mathbb{Z}^n$ in $C(2B)$, the resulting $N(B, k)$ hypercubes will not intersect, but they will all be contained within the enlarged hypercube $C(2B + \lambda)$, thus giving $N(B, k) \leq \frac{\mathrm{vol}(C(2B+\lambda))}{\mathrm{vol}(C(\lambda))} = (\frac{2B}{\lambda} + 1)^n$. To derive a lower bound on λ, note that for any $\mathbf{t} \in I_{S'}$ we have that the norm $\mathcal{N}(\mathbf{t}) = \mathcal{N}(\langle t \rangle) \geq \mathcal{N}(\langle \Phi_{S'}, q \rangle) = q^{kd_q}$, where the inequality holds because $\langle t \rangle \subseteq \langle \Phi_{S'}, q \rangle$ as an ideal, and the last equality is because $\deg \Phi_{S'} = kd_q$.

It follows from the arithmetic-geometric inequality that $\|\mathbf{t}\| = \frac{1}{\sqrt{n}} T_2(\mathbf{t}) \geq \mathcal{N}(\mathbf{t})^{1/n} \geq q^{k/k_q}$. By equivalence of norms, we can conlude that $\|\mathbf{t}\|_\infty \geq \lambda \geq \frac{1}{\sqrt{n}} q^{k/k_q}$. Finally, using $B = \frac{1}{\sqrt{n}} q^\beta$, for $k \geq \beta k_q$, we have $\lambda \geq B$, so $N(B, k) = 0$, while for $k < \beta k_q$, we have

$$N(B, k) \leq \left(\frac{2B}{\lambda} + 1 \right)^n \leq (2q^{\beta - k/k_q} + 1)^n \leq 2^{2n} q^{(\beta - k/k_q)n}$$

as claimed.

For the second bound we claim that $N(B, k) \leq (2B)^{n - kd_q} = (2B)^{n(1 - k/k_q)}$. In fact, the degree of $\Phi_{S'}$ is kd_q, so the vector $\bar{\mathbf{t}}$ formed by the $n - kd_q$ low-order coefficients of $\mathbf{t} = \Phi_{S'} \mathbf{t}'$ is related to the vector $\mathbf{t}'$ formed by the $n - kd_q$ low-order coefficients of $\mathbf{t}'$ by a lower triangular $(n - kd_q) \times (n - kd_q)$ matrix whose diagonal elements are the non-zero constant coefficients a of $\Phi_{S'}$. That means this matrix is non-singular modulo q and the mapping from $\mathbf{t}'$ to $\mathbf{t}$ is one-to-one providing the claim.

Since the number of $s \in R_q/I_S$ divisible by $\Phi_{S'}$ is $q^{d_q(|S|-k)}$, the upper bound given above implies that

$$p \leq 2^{(m+1)|S|} \max_{0 \leq k \leq |S|} \frac{N(B,k)^m}{q^{(m-1)(|S|-k)d_q}}.$$

Using the first bound we gave on $N(B,k)$ with $B = \frac{1}{\sqrt{n}} q^{\beta}$ we get

$$p \leq 2^{(m+1)(|S|+2n)} \max_{0 \leq k < \beta k_q} q^{n\left(m(\beta - \frac{k}{k_q}) - (m-1)\frac{|S|-k}{k_q}\right)}.$$

Viewing the exponent on the right hand side as a function of k, it is clear that it reaches its maximum for $k = 0$, assuming the value $-mn\varepsilon$ when $\beta = (1 - \frac{1}{m})\frac{|S|}{k_q} - \varepsilon$ and giving therefore the first wanted bound. If $|S| = k_q$ we can use the second bound we gave on $N(B,k)$ with $B = q^{\beta}$. Since in this case $N(B,k) = 0$, we have

$$p \leq 2^{(m+1)(|S|+2n)} \max_{0 \leq k < \beta k_q} q^{n((1-\beta)m-1)\left(\frac{k}{k_q} - 1\right)} = 2^{(m+1)(|S|+2n)} q^{-\frac{n}{k_q}((1-\beta)m-1)}$$

where the last equality holds for any $\beta \leq 1 - \frac{1}{m}$. Using $\beta = 1 - \frac{1}{m} - k_q\varepsilon$, we obtain the wanted result on $\lambda_1^{\infty}(L(\mathbf{a}, I_S))$. □

In our analysis of the distribution of the NTRU key g/f with $k_q = O(1)$, we will also use the following lower bound on $\lambda_1(\mathbf{a}^{\perp}(I_S))$.

Lemma 1.5.3. *Let $n \geq 8$ be a power of 2 and $q \geq 5$. Assume that $\Phi = x^n + 1$ splits into k_q distint irreduible factors modulo q of degree $d_q = n/k_q$. Then, for $m \geq 2$ and $\varepsilon > 0$, we have*

$$\lambda_1^{\infty}(\mathbf{a}^{\perp}(I_S)) \geq \begin{cases} \frac{1}{\sqrt{n}} q^{\frac{1}{m} + (m-1)\frac{|S|}{d_q} - \varepsilon} & \text{for any } 0 \leq |S| \leq k_q \\ q^{\frac{1}{m} - k_q\varepsilon} & \text{for } |S| = 0 \end{cases}$$

with probability greater than $1 - 2^{4mn} q^{-\varepsilon mn}$ over the uniformly random choice of $\mathbf{a} \in (R_q^{\times})^m$.

Proof. Let p denote the probability over $\mathbf{a}$ that $L(\mathbf{a}^{\perp}(I_S))$ contains a non-zero vector $\mathbf{t}$ of infinity norm less than B. We can bound p from above summing the probabilities $p(\mathbf{t}) = \Pr_{\mathbf{a}}[\sum_{i \leq m} a_i t_i = 0 \mod q]$ over all possible values for $\mathbf{t}$ with $\|\mathbf{t}\|_{\infty} < B$ and $t_i \in I_S$ for $i = 1, \ldots, m$.

By the Chinese Remainder Theorem, we have $p(\mathbf{t}) = \prod_{j \leq k_q} p_j(\mathbf{t})$, where $p_j(\mathbf{t}) = \Pr_a[\sum_{i \leq m} a_i t_i = 0 \mod \Phi_j]$. Let $\Phi_S = \prod_{i \in S} \Phi_i$, $\Phi_{\overline{S}} = \prod_{i \in \overline{S}} \Phi_i$ and $\Phi_{S'} = \gcd(t_1, \ldots, t_m, \Phi_{\overline{S}}) = \prod_{i \in S'} \Phi_i$ for some $S' \subseteq \overline{S}$ of cardinality $0 \leq k \leq |\overline{S}|$. For any $j \in S \cup S'$, we have $\sum_{i \leq m} t_i a_i = 0 \mod \Phi_j$ regardless of the value of $a_i \mod \Phi_j$. On the other hand, or any $j \in \overline{S} \setminus S'$, there exists $i \leq m$ such that $t_i \neq 0 \mod \Phi_j$ so that for any choice of $\{a_j\}_{j \neq i}$, there is a unique value of $a_i \mod \Phi_j$ such that $\sum_{i \leq m} t_i a_i = 0 \mod \Phi_j$. From this follows that $p_j(\mathbf{t}) = \frac{1}{q^{d_q - 1}}$. As a consequence, we have $p(\mathbf{t}) = \frac{1}{(q^{d_q}-1)^{|\overline{S}|-k}}$ and

$$p \leq \sum_{0 \leq k \leq |\overline{S}|} \sum_{\substack{S' \subseteq \overline{S} \\ |S'|=k}} \sum_{\substack{\mathbf{t} \in R_q^m \\ \forall i, 0 < \|ti\|_\infty < B \\ \forall i, \Phi_S \Phi_{S'} | t_i}} \frac{1}{(q^{d_q} - 1)^{|\overline{S}|-k}}.$$

For S' with $|S'| = k$, let $N(B,k)$ denote the number of $t \in R_q$ such that $\|t\|_\infty < B$ and $t = \Phi_S \Phi_{S'} t'$ for some $t' \in R_q$ of degree less than $n(1-(k+|S|)/k_q)$. Like for the previous lemma, we derive two upper bounds for $N(B,k)$, from which we get the wanted bounds on $\lambda_1^\infty(L(\mathbf{a}, I_S))$. The first upper bound, with $B = \frac{1}{\sqrt{n}} q^\beta$, shows that $N(B,k) = 0$ for $k \geq \beta k_q - |S|$, while $N(B,k) \leq 2^{2n} q^{(\beta-(|S|+k)/k_q)n}$ for $k < \beta k_q - |S|$. The second bound is $N(B,k) \leq (2B)^{n(1-(|S|+k)/k_q)}$. The first bound on $N(B,k)$ with $B = \frac{1}{\sqrt{n}} q^\beta$, leads to

$$p \leq 2^{2|\overline{S}|+2n} \max_{0 \leq k \leq k_q} q^{n\left(m(\beta - \frac{|S|+k}{k_q}) - \frac{|\overline{S}|-k}{k_q}\right)}.$$

Viewing again the right hand side exponent as a function of k, it is maximized for $k = 0$ with value $-mn\varepsilon$ when $\beta = \frac{1}{m} + (1 - \frac{1}{m})\frac{|S|}{k_q} - \varepsilon$ giving the first bound.

In the case $|S| = 0$, using our second bound on $N(B,k)$ with $B = q^\beta$ and noting that $N(B,kq) = 0$, we get

$$p \leq 2^{2|\overline{S}|+n} \max_{0 \leq k < k_q} q^{n(1-m\beta)\left(\frac{k}{k_q} - 1\right)} = 2^{2|\overline{S}|+n} q^{n(1-m\beta)\left(1 - \frac{1}{k_q}\right)}$$

where the last equality holds for any $\beta \leq \frac{1}{m}$. Using $\beta = \frac{1}{m} - k_q \varepsilon$ we have the seond claimed bound. □

1.5.2 Regularity bounds for ring R_q

In this section we want to study the closeness to uniformity of the distribution of $(m+1)$-tuples from $(R_q^\times)^m \times R_q$ of the form $(a_1, \ldots, a_m, \sum_{i \leq m} t_i a_i)$, where the a_i's are independent and uniformly random in $R_q^\times$, and the t_i's are chosen from some distribution on R_q concentrated on elements of small height. Similarly to [Mic07], we call *regularity* of the generalized knapsak funtion $(t_i)_{i \leq m} \mapsto \sum_{i \leq m} t_i a_i$ the statistical distance of such distribution to the uniform one on $(R_q^\times)^m \times R_q$. In our particular case, for NTRU applications we are interested in the case where $m = 2$.

The result in [Mic07] yelds that when the a_i's are uniformly random in the whole ring R_q and the t_i's are uniformly random on the subset of elements of R_q of height at most d (for some $d < q$), the regularity bound is $\Omega(\sqrt{nq/d^m})$. Unfortunately, for small values of m and q (as in our case $m = O(1)$ and $q = poly(n)$) this bound is non-negligible. In order to make it exponentially small in n one needs to set $m \log(d) = \Omega(n)$, which in turn would lead to inefficient cryptographic functions. Now R_q contains n proper ideals of size $q^{n-1} = |R_q|/q$, and the probability $\approx n/q^m$ that all of the a_i's fall into one such ideal (which implies $\sum t_i a_i$ to also be in the proper ideal) is non-negligible for small m. This means that when the a_i's are chosen uniformly from the whole ring R_q with $q = 1 \mod 2n$, the effective regularity bound we have to deal with is not much better than the one given above. To avoid this problem, we will restrict the a_i's to be uniform in $R_q^\times$ and choose the t_i's from a discrete Gaussian distribution, using a different argument to prove an exponentially small bound in n.

As a direct consequence of Lemmata 1.2.9, 1.2.14, 1.5.1 and 1.5.2, we have the following:

Lemma 1.5.4. *Let $n \geq 8$ be a power of 2 such that $\Phi = x^n + 1$ splits into k_q irreducible factors modulo a prime $q \geq 5$. Let $S \subseteq \{1, \ldots, k_q\}$, $m \geq 2$, $\varepsilon > 0$, $\delta \in (0, 1/2)$ and $\mathbf{t} \sim D_{\mathbb{Z}^{mn},\sigma}$, $\mathbf{c} \in \mathbb{Z}^{mn}$ with*

$$\sigma \geq \begin{cases} \sqrt{n \log(2mn(1 + 1/\delta))/\pi}\, q^{1-(1-\frac{1}{m})(1-\frac{|S|}{k_q})+\varepsilon} & \text{for any } 0 \leq |S| \leq k_q \\ \sqrt{n \log(2mn(1 + 1/\delta))/\pi}\, q^{\frac{1}{m}+k_q\varepsilon} & \text{for } |S| = 0. \end{cases}$$

Then for all but a fraction at most $2^{4mn} q^{-\varepsilon mn}$ of $\mathbf{a} \in (R_q^\times)^m$ we have

$$\Delta\Big[\mathbf{t} \mod \mathbf{a}^\perp(I_S), U(R/\mathbf{a}^\perp(I_S))\Big] < 2\delta.$$

Using the previous result, we can finally prove our bound.

Theorem 1.5.5. *Let $n \geq 8$ be a power of 2 such that $\Phi = x^n + 1$ splits into k_q irreducible factors modulo a prime $q \geq 5$. Let $m \geq 2$, $\varepsilon > 0$, $\delta \in (0, 1/2)$ and $\mathbf{t} \sim D_{\mathbb{Z}^{mn},\sigma}$, with $\sigma \geq \sqrt{\log(2mn(1 + 1/\delta))/\pi} \min(\sqrt{n} q^{\frac{1}{m}+\varepsilon}, q^{\frac{1}{m}+k_q\varepsilon})$. Then for all except a fraction at most $2^{4mn} q^{-\varepsilon mn}$ of $\mathbf{a} \in (R_q^\times)^m$, we have $\eta_\delta(\mathbf{a}^\perp) \leq \sqrt{\log(2mn(1 + 1/\delta))/\pi} \min(\sqrt{n} q^{\frac{1}{m}+\varepsilon}, q^{\frac{1}{m}+k_q\varepsilon})$, and the distance to uniformity of $\sum_{i \leq m} t_i a_i$ is at most 2δ. As a consequence*

$$\Delta\left[\Big(a_1, \ldots, a_m, \sum_{i \leq m} t_i a_i\Big), U\Big((R_q^\times)^m \times R_q\Big)\right] \leq 2\delta + 2^{4mn} q^{-\varepsilon mn}.$$

Proof. For each $\mathbf{a} \in (R_q^\times)^m$, let $D_{\mathbf{a}}$ denote the distribution of $\sum_{i \leq m} t_i a_i$ where $\mathbf{t}$ is sampled from $D_{\mathbb{Z}^{mn},\sigma}$. If we call $\Delta_{\mathbf{a}}$ the distance to uniformity of $D_{\mathbf{a}}$, the above statistical distance is exactly $\frac{1}{|R_q^\times|^m} \sum_{\mathbf{a} \in (R_q^\times)^m} \Delta_{\mathbf{a}}$. To prove the theorem it is then enough to show a uniform bound $\Delta_{\mathbf{a}} \leq 2\delta$, for all except a fraction of at most $2^{4mn} q^{-\varepsilon mn}$ of the possible $\mathbf{a} \in (R_q^\times)^m$. Now, the mapping

$\phi : \mathbf{t} \mapsto \sum_i t_i a_i$ induces an isomorphism from the quotient group $\mathbb{Z}^m n/\mathbf{a}^\perp$ to R_q, which is the image of ϕ for to the invertibility of the a_i's. Hence, the statistical distance $\Delta_{\mathbf{a}}$ is equal to the distance to uniformity of $\mathbf{t} \mod \mathbf{a}^\perp$. Since it is needed to study the NTRU key generation algorithm, we will also study the distance to uniformity of $\mathbf{t} \mod \mathbf{a}^\perp(I_S)$ for any $S \subseteq \{1, \ldots, k_q\}$. By Lemma 1.2.14, we already have $\Delta_{\mathbf{a}} \leq 2\delta$ if $\sigma > \eta_\delta(\mathbf{a}^\perp(I_S))$. Applying Lemma 1.2.9 we can bound $\eta\delta(\mathbf{a}^\perp(I_S))$ from above, reducing ourself to bound from below the minimum of the dual lattice. By Lemma 1.5.1 we know said lattice to be $\mathbf{a}^\perp(I_S)^\vee = \frac{1}{q} L(\mathbf{a}^\star, I^\star_{\overline{S}})$, and we know how to do this by Lemma 1.5.2. To conclude, it is enough to use Lemma 1.5.4 with $S = \emptyset$ and $\mathbf{c} = \mathbf{0}$. □

Chapter 2

Modern lattice problems

In [Reg09], Regev introduced the average-case problem called Learning with Errors Problem (LWE). Since then it has appeared as the most suitable lattice problem to support an encryption scheme and it has also enabled to build a chosen ciphertext-secure cryptosystem [PW11] and identity based encryption schemes [CHKP12, GPV08].

We will here give some details and properties on LWE, then we discuss its hardness and efficiency. Unfortunately, a loss in efficiency was necessary to prove the hardness in the classical case, making LWE-based protocols impractical for real-life use. For this reason in [LPR12] Lyubashevsky, Micciancio and Regev added algebraic structure (namely a ring structure) to LWE, and this is what we will study in the second part of this chapter.

2.1 Learning With Errors Problem (LWE)

Let n, q and m be integers, ψ an error distribution over $\mathbb{Z}$.

Definition 2.1.1. For a vector $\mathbf{s} \in \mathbb{Z}_q^n$, called *secret*, the *LWE distribution* $A_{s,\psi}$ over $\mathbb{Z}_q^n \times \mathbb{Z}_q$ is sampled by choosing $\mathbf{a} \in \mathbb{Z}_q^n$ uniformly at random, choosing $e \sim \psi$, and giving in output $(\mathbf{a}, b = \langle \mathbf{s}, \mathbf{a} \rangle + e \mod q) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$.

In practice ψ can be thought as a discrete Gaussian of width αq for some $\alpha < 1$, *i.e.* $\psi = D_{\mathbb{Z},\alpha q}$.

Definition 2.1.2. Given m independent samples $\{(\mathbf{a}_i, b_i)\}_{i=1}^m \in \mathbb{Z}_q^n \times \mathbb{Z}_q$ drawn from the LWE distribution $A_{s,\psi}$, with $\mathbf{s}$ chosen uniformly at random, the *search* version of the *Learning With Error Problem* consists in finding $\mathbf{s}$.

An intuitive way to see the problem is to consider it as trying to find the solution $\mathbf{s} \in \mathbb{Z}_q^n$ of the system of linear equations with "errors":

$$\begin{cases} \langle \mathbf{s}, \mathbf{a}_1 \rangle = b_1 - e_1 \pmod q \\ \langle \mathbf{s}, \mathbf{a}_2 \rangle = b_2 - e_2 \pmod q \\ \vdots \\ \langle \mathbf{s}, \mathbf{a}_n \rangle = b_n - e_n \pmod q, \end{cases}$$

where each $\mathbf{a}_i$ is uniformly random and $e_i \sim \psi$ for $i = 1, \dots, n$. The same problem can be given a compact matrix expression

$$\mathbf{b}^T = \mathbf{s}^T A + \mathbf{e}^T \mod q.$$

Definition 2.1.3. Given an error distribution ψ over $\mathbb{Z}$ and m independent samples $\{(\mathbf{a}_i, b_i)\}_{i=1}^m \in \mathbb{Z}_q^n \times \mathbb{Z}_q$, where every sample is either drawn according to $A_{\mathbf{s},\psi}$ for a fixed and uniformly random $\mathbf{s} \in \mathbb{Z}_q^n$, or the uniform distribution, distinguish which is the case.

As suggested by the matrix expression, Search-LWE can be equivalently presented as an average case of BDD$_\gamma$ over the q-ary lattices $L_q(A) = \{\mathbf{y} \in \mathbb{Z}^m \mid \mathbf{y} = A^T\mathbf{s} \mod q \text{ for some } \mathbf{s} \in \mathbb{Z}^n\}$. In this setting, the vector $\mathbf{b}$ is relatively close only to one vector $L_q(A)$ and we can notice that LWE problem consists in finding this “target”. In [Reg09] Regev proved the following worst-case to average-case reduction involving the decision version of LWE :

Theorem 2.1.4. *For any $m = poly(n)$, $q \leq 2^{poly(n)}$ and any discrete Gaussian error distribution ψ of parameter $\alpha q \geq 2\sqrt{n}$, with $0 < \alpha < 1$, solving the Decision-LWE with parameters n, q, ψ, m is at least as hard as solving GapSVP$_\gamma$ and SIVP$_\gamma$ on arbitrary $n-$dimensional lattices, for some $\gamma = \widetilde{O}(n/\alpha)$, using quantum computation.*

In [Pei09], Peikert managed to make the above reduction completely classical, but only with the following conditions:

1. the classical reduction only involves GapSVP$_\gamma$, while the quantum works also for SIVP$_\gamma$;

2. q has to be exponentially large, more precisely $q \geq 2^{n/2}$, as opposed to $q \geq 2^{\sqrt{n}/\alpha}$ with $0 < \alpha < 1$.

Such an undesirable bound on q forces the key size to be larger, and therefore leads to less efficiency for all cryptographic protocols based on this problem. Nevertheless, the techniques applied by Peikert were used by Lyubashevski and Micciancio in [LM09] to prove that, for $\gamma = poly(n)$, GapSVP$_\gamma$, SVP$_\gamma$ and BDD$_\gamma$ are equivalent problems.

2.2 Learning With Errors Problem over rings (R-LWE)

To improve efficiency of LWE-based protocols, Lyubashevsky, Peikert and Regev introduced the ring-learning with errors problem [LPR12], an analogue problem of the LWE in the ring setting, whose hardness can be linked to some worst-case problem over ideal lattices. In this section we will introduce this problem and give some ideas on the worst-case to average-case reduction as in [LPR12, LPR13].

2.2.1 Error distributions

Here we the introduce the family of error distributions we need using in order to define the problem and for which the worst-case to average-case reduction effectively works.

Definition 2.2.1. Let $\alpha > 0$ be a real number, the family of error distributions $\Psi_{\leq \alpha}$ is the set of all elliptical Gaussian distributions D_σ over $K_\mathbb{R}$, where, for any parameter $\sigma = (\sigma_1, \dots, \sigma_n)$, we have $\sigma_i \leq \alpha$.

Definition 2.2.2. The gamma distribution $\Gamma(2,1)$ with shape parameter 2 and scale parameter 1 is the distribution with the following density:

$$f(x) = \begin{cases} xe^{-x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0. \end{cases}$$

Definition 2.2.3. Let α be a positive real number, then a distribution sampled from Υ_α is an elliptical Gaussian distribution D_σ over $K_\mathbb{R}$ whose parameters $r_i > 0$ are such that $\sigma_i^2 = \sigma_{i+n/2}^2 = \alpha^2(1 + \sqrt{n}x_i)$, with $x_1, \dots, x_n \in \mathbb{R}$ chosen independently from $\Gamma(2,1)$.

2.2.2 The general instance

The parameters of Ring-LWE are a number field K with $R = \mathcal{O}_K$ and a prime $q \geq 2$. We here define the problem in the most general case in which K is a number field, even though the worst-case to average-case reduction for R-LWE has been proved only for cyclotomic fields.

Let us denote $\mathbb{T} = K_\mathbb{R}/R^\vee$ (remark that $R^\vee$ is seen as the codifferent ideal).

Definition 2.2.4. Let $s \in R_q^\vee$ be the secret and ψ an error distribution over $K_\mathbb{R}$, then a sample from the ring-LWE distribution $A_{s,\psi}$ over $R_q \times \mathbb{T}$ is generated by choosing $a \sim U(R_q)$, $e \sim \psi$ and giving the output $(a, b = (as)/q + e \mod R^\vee)$.

Definition 2.2.5 (Search R-LWE$_{s,\Psi}$)**.** Let Ψ be a family of distributions over $K_{\mathbb{R}}$. Given arbitrarily many independent samples from $A_{s,\psi}$, for some $s \in R_q^\vee$ and $\psi \in \Psi$, the search version of ring-LWE is the problem of finding s.

Definition 2.2.6 (Decision, R-DLWE$_{s,\Upsilon}$)**.** Let Υ be a distribution over a family of error distribution over $K_{\mathbb{R}}$. The *average-case decision ring-LWE problem* consists in distinguishing with non-negligible advantage between arbitrarily many independent samples from $A_{s,\psi}$, with $(s, \psi) \sim U(R_q^\vee) \times \Upsilon$, and the same number of uniformly random and independent samples from $R_q \times \mathbb{T}$.

2.2.3 Hardness

Just as for the LWE, the reduction from the R-LWE to some worst-case lattice problem involves quantum computing. Here we give the statement of the hardness problem and the idea of the proof.

Theorem 2.2.7. *Let K be the m-th cyclotomic number field, of dimension $n = \varphi(m)$ and let $R = \mathcal{O}_K$ be its ring of integers. Let $\alpha = \alpha(n) > 0$ and let $q = q(n) \leq poly(n)$, $q = 1 \mod m$ be a prime such that $\alpha q \geq \omega(\sqrt{\log n})$. Then there is a polynomial time quantum reduction from approximate $SIVP_\gamma$ with $\gamma = \widetilde{O}(\sqrt{n}/\alpha)$ over ideal lattices to the decision R-LWE$_{q,\Upsilon_\alpha}$.*

The proof composes of two (basically independent) parts.

Part I: Worst-case hardness of the search problem

For an opportune choice of parameters, the R-LWE$_{q,\psi}$ is at least as hard as quantumly solving SIVP$_\gamma$ on ideal lattices of $\mathbb{R}$. It is important to notice that this reduction actually works in any number field, not only for cyclotomic ones. More precisely we have the following result.

Theorem 2.2.8. *Let K be an arbitrary number field of degree n, $R = \mathcal{O}_K$, $\alpha = \alpha(n) > 0$ and $q = q(n) \geq 2$ be such that $\alpha q \geq \omega(\sqrt{\log n})$. Then there is a probabilistic polynomial time quantum reduction from approximate $SIVP_\gamma$, with $\gamma = \widetilde{O}(\sqrt{n}/\alpha)$, to R-LWE$_{q,\Psi_{\leq\alpha}}$.*

The proof of this result follows the line of [Reg09] for general lattices, applying repeatedly an iterative step with the goal of finding shorter and shorter vectors. There is to notice that, as of the time of this work, no way has been found to replace the use of quantum computing in the proof.

Part II: Decision-to-Search reduction

The second part consists in showing that solving the decision version of R-LWE is at least as hard as solving its search variant. This implies that if SIVP$_\gamma$ is hard to solve in the quantum setting, then the Ring-LWE

Distribution is pseudorandom. Such reduction is entirely classical and it relies on cyclotomic number fields being Galois fields and on the particular choice of the modulus q, such that $q\mathcal{O}_K$ splits completely into n prime ideals $\mathfrak{q}_i$.

The formal result is the following:

Theorem 2.2.9. *Let R and q be as above and let $\alpha q \geq \eta_\varepsilon(R^\vee)$ for some negligible $\varepsilon = \varepsilon(n)$. Then there is a randomized polynomial time reduction from R-LWE$_{q,\Psi_\alpha}$ to R-DLWE$_{q,\Upsilon_\alpha}$.*

The proof is composed as the concatenation of four different reductions, which we will now present briefly. Full details can be found in [LPR12].

$$\text{R-LWE}_{q,\Psi} \xrightarrow{(1)} \mathfrak{q}_i\text{-LWE} \xrightarrow{(2)} WDLWE^i_{q,\Psi} \xrightarrow{(3)} DLWE^i_{q,\Upsilon} \xrightarrow{(4)} DLWE_{q,\Upsilon}$$

(1) (R-LWE$_{q,\Psi}$ to $\mathfrak{q}_i$-LWE) Given access to $A_{s,\psi}$ for some arbitrary $s \in R_q^\vee$ and $\psi \in \Psi_{\leq\alpha}$ for some $\alpha > 0$, we call $\mathfrak{q}_i$-LWE the problem of finding s mod $\mathfrak{q}_i R^\vee$. This can be seen as a local variant of the general problem, and the reason why this reduction works is the fact that the Galois group acts transitively on the $\mathfrak{q}_i$'s.

Lemma 2.2.10. *For every $i \in \mathbb{Z}_m^\times$ there is a deterministic polynomial-time reduction from R-LWE$_{q,\Psi_\alpha}$ to $\mathfrak{q}_i$-LWE$_{q,\Psi_\alpha}$.*

(2) ($\mathfrak{q}_i$-LWE to WDLWE$^i_{q,\Psi}$) For $i \in \mathbb{Z}_m^\times$ and a family of distributions Ψ and given access to $A^j_{s,\psi}$ for arbitrary $s \in R_q^\vee$, $\psi \in \Psi$ and $j \in \{i-, i\}$, the worst-case decision $\mathfrak{q}_i$-LWE is the problem of finding j and it is denoted as WDLWE$^i_{q,\Psi}$.

Lemma 2.2.11. *For any $i \in \mathbb{Z}_m^\times$ there is a probabilistic polynomial time reduction from $\mathfrak{q}_i$-LWE to WDLWE$^i_{q,\Psi}$.*

(3) (WDLWE$^i_{q,\Psi}$ to DLWE$^i_{q,\Upsilon}$) Since we start from WDLWE$^i_{q,\Psi}$ being a local worst-case problem, we want move to an average-case local problem first.

For any $i \in \mathbb{Z}_m^\times$ and Υ distribution over error distributions, the average-case decision $\mathfrak{q}_i$-LWE is denoted as DLWE$^i_{q,\Upsilon}$ and it is the problem to distinguish, over random choices $(s, \psi) \sim U(R_q^\vee) \times \Upsilon$ and with non-negligible advantage, between inputs from $A^i_{s,\psi}$ versus inputs from $A^{i-}_{s,\psi}$.

Lemma 2.2.12. *For any $\alpha > 0$ and any $i \in \mathbb{Z}_m^\times$, there is a randomized polynomial time reduction from WDLWE$^i_{q,\Psi_{\leq\alpha}}$ to DLWE$^i_{q,\Upsilon_\alpha}$.*

(4) (DLWE$^i_{q,\Upsilon}$ to DLWE$_{q,\Upsilon}$) The last reduction finally removes the dependence on a specific $\mathfrak{q}_i$.

Lemma 2.2.13. *Let Υ be a distribution over a family of error distributions such that for any $\psi \in \Upsilon$ and any $s \in R_q^\vee$ the distribution $A_{s,\psi}^{m-1}$ is within negligible statistical distance from the uniform. Then for any oracle solving the DLWE$_{q,\Upsilon}$ problem, there exist an $i \in \mathbb{Z}_m^\times$ and an efficient algorithm that solves DLWE$^i_{q,\Upsilon}$ using this oracle.*

2.2.4 Variants of R-LWE

For our purposes, we want to use the polynomial representation rather than the one just defined, and have a discrete noise distribution. Therefore we need to rephrase the definition of the problem and adjust the noise distribution to better suit our necessities.

Definition 2.2.14. Let Υ be a distribution over a family of distributions on R. The *Ring Learning With Errors Problem* with parameters q and Υ R-LWE$_{q,\Upsilon}$ is as follows: let ψ be sampled from Υ and $\mathbf{s}$ be chosen uniformly in R_q. Given access to an oracle O that produces samples in $R_q \times R_q$, distinguish whether O outputs samples from the distribution $A_{\mathbf{s},\psi}$ or $U(R_q \times R_q)$ with non-negligible advantage.

Again, R-LWE an be interpreted as a problem over module q-ary lattices. Let m be the number of samples asked to the oracle, and let $\{(\mathbf{a}_i, b_i)\}_{i=1}^m$ be the samples. Then solving R-LWE consists in distinguishing whether the vector $\mathbf{b}$ is generated uniformly modulo the (module) lattice $L_q(a)$ or around the origin according to some Gaussian-like distribution and then reduced modulo the lattice. We can also adapt a result in [LPR12] in the following form:

Theorem 2.2.15. *Assume $\alpha q = \omega(n\sqrt{\log n})$ with $\alpha \in (0,1)$ and $q = poly(n)$ prime with $q = 1 \mod 2n$. Consider then the distribution $\overline{\Upsilon}_\alpha$. There exists a randomized polynomial-time quantum reduction from Ideal-SVP$_\gamma$ to R-LWE$_{q,\Upsilon_\alpha}$ (denoted by R-LWE$_{q,\alpha}$ in the sequel), with $\gamma = \omega(n^{\frac{3}{2}} \log n)/\alpha$.*

For $s \in R_q$ and ψ a distribution in R_q, we denote as $A_{s,\psi}^\times$ the distribution obtained by sampling the pairs $(a, as+e)$ with $a \sim R_q^\times$ and e sampled independently from ψ. When $q = \Omega(n)$, the probability for a uniform element of R_q of being invertible is non-negligible, so R-LWE remains hard even when $A_{s,\psi}$ and $U(R_q \times R_q)$ are respectively replaced by $A_{s,\psi}^\times$ and $U(R_q^\times \times R_q)$. Moreover, we can add the secret s to be chosen from the error distribution without any security reduction. We will refer to this variant as R-LWE$_{\text{HNF}}^\times$.

We need to define what is $\overline{\Upsilon}_\alpha$, and that's what we will do now. For $\sigma \in (\mathbb{R}^+)^n$, we define the elliptical Gaussian ρ_σ like before as the row vector of independent Gaussians $(\rho_{\sigma_1}, \dots, \rho_{\sigma_n})$, where $\sigma_i = \sigma_{i+n/2}$ for $1 \le i \le n/2$. At this point, since we want to use the polynomial representation rather than the space H, we multiply ρ_σ from the right first by $M = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \otimes Id_{n/2} \in$

$\mathbb{C}^{n\times n}$, and then $V \in C^{n\times n}$, where V is the matrix whose upper half is $\frac{1}{n}(\zeta^{-(2j+1)k})_{0\leq j<n/2,0\leq k<n}$ and the bottom half is the complex conjugate of the upper half. We will call the result ρ'_σ. We then define a sample from $\overline{\rho}'_\sigma$ computing a sample from ρ'_σ with the absolute error which is at most $1/n^2$. If said sample is within distance $1/n^2$ of a half integer, we resample, otherwise we round it to the closest integer and reduce it modulo q. Finally, a distribution sampled from $\overline{\Upsilon}_\alpha$ for $\alpha \geq 0$ is defined as $\overline{\rho}'_\sigma$, where $\sigma_i^2 = \sigma_{i+n/2}^2 = \alpha^2 q^2(1+\sqrt{n}x_i)$, with the x_i's sampled independently from the distribution $\Gamma(2,1)$ for $i \leq n/2$.

Remark 2.2.16. This definition ends up being very close to the original one in Section 2.2.2, but with the fundamental difference that the sampling now uses a rejection process to round to R. Nonetheless, the problem remains hard because samples pass the rejection step with non-negligible probability, and the rounding can be performed on the oracle samples without considering the actual error.

All sampling from these distributions can be computed in quasi-polynomial time, moreover samples from $\overline{\Upsilon}_\alpha$ are very small.

Lemma 2.2.17. *Let $y, r \in R$, y be sampled accordingly to $\overline{\Upsilon}_\alpha$, with $\alpha q \geq n^{1/4}$. Then*

(*a*) $\Pr\left[\|yr\| \geq \alpha q n^{1/4}\omega(\sqrt{\log n})\|r\|\right] \leq n^{-\omega(1)}$*;*

(*b*) $\Pr\left[\|yr\|_\infty \geq \alpha q n^{-1/4}\omega(\log n)\|r\|\right] \leq n^{-\omega(1)}$.

Proof. Define Υ_α as $\overline{\Upsilon}_\alpha$ without the rejection step, because of the bound on the rejection probability, it is enough to show the result with Υ_α. Let $(r^{(k)})_k$ be the embedding vector of r. Multiplying y by r is the same as sampling from $\rho_{\sigma'}$ with $\sigma' k = \sigma'_{k+n/2} = \sigma_k|r^{(k)}|$ (see [LS15] for a proof), thus we have $\sigma'_k \leq \alpha q n^{1/4}\omega(\sqrt{\log n})|r^{(k)}|$ for all $k \leq n$, with a probability which is at least $1 - n^{-\omega(1)}$.

To obtain the coefficients of yr, we apply M and V to the vector of the samples. The magnitude of the entries of the product matrix is at most $O(1/n)$, so the coefficients of the polynomial yr are distributed as statistically independent (one-dimensional) Gaussians of standard deviations at most $\alpha q n^{-1/4}\omega(\sqrt{\log n})\|r\|$, implying the Euclidean norm of the n-dimensional vector to be at most $\alpha q n^{1/4}\omega(\sqrt{\log n})\|r\|$ with probability greater than $1 - n^{\omega(1)}$.

Now all the coordinates are bounded by $\alpha q n^{-1/4}\omega(\log n)\|r\|$ with probability at least $1 - n^{-\omega(1)}$. The additional rounding error $O(\sqrt{n})$ only changes the hidden constant factor in the $\omega(\log n)$ factor, thanks to the hypothesis of $\alpha q \geq n^{1/4}$, so the proof is complete. □

Chapter 3

A provably secure variant of NTRU cryptosystem

As seen in section 1.4, the public key h for an NTRU instance is the ratio of two randomly generated polynomials f and g both with small coefficients. Our goal is to modify the original scheme in order to derive IND-CPA (Indistinguishability under Chosen-Plaintext Attack)[1] security from R-LWE by making sure the distribution of h is statistically very close to the uniform distribution over $R_q^\times$.

Doing so we will provide new key generation algorithms as well as a new NTRUEncrypt scheme.

3.1 A revised key generation algorithm for NTRU-Encrypt

We now want to use the results of the previous chapter to derive key generation algorithms for the NTRU scheme, to be able to generate public keys following the distributions for which Ideal-SVP is known to reduce to R-LWE.

The secret key polynomials f and g will be generated using the Gentry et al. sampler of Lemma 1.2.15, and rejected until the output polynomials are invertible modulo q. This sampler may not exactly sample from discrete Gaussians, but since the statistical distance can be made negligible, the impact on our results is also negligible. Moreover the conditions we will use on standard deviations are much stronger than the one in Lemma 1.2.15. That being said, from now on we will assume we have a perfect discrete Gaussian sampler.

[1] Suppose to have two plaintext messages M_0, M_1, and a bit b, and let M_b be encrypted to a ciphertext C. Having IND-CPA means that no method for recovering b from C without knowledge on the secret key has a success probability non-negligibly greater than 1/2.

In the notation above, the algorithm is the following:

```
Algorithm 3 Encryption Key Generation
Input: $n, q \in \mathbb{Z}, p \in R_q^\times, \sigma > 0$
1: Sample $f'$ from $D_{\mathbb{Z}^n,\sigma}$ and compute $f = pf' + 1$
2: if $(f \mod q) \notin R_q^\times$ then
3:     resample
4: Sample $g$ from $D_{\mathbb{Z}^n,\sigma}$
5: if $(g \mod q) \notin R_q^\times$ then
6:     resample
7: Return secret key $sk = f$ and public key $pk = h = pg/f \in R_q^\times$
Output: The key pair $(sk, pk) \in R \times R_q^\times$
```

Remark 3.1.1. By choosing a large enough standard deviation σ, we can apply the results of Section 1.5 and obtain the (quasi-)uniformity of the public key. We sample f of the form $pf' + 1$ so that it has inverse 1 modulo p, making the decryption process of NTRUEnrypt more efficient (as in the original NTRUEnrypt scheme). There is to note that the rejection condition on f at Step 1 is equivalent to the $(f' \mod q) \notin R_q^\times - p^{-1}$, where p^{-1} is the inverse of p in R_q.

The following result ensures that, for an appropriate choice of the parameters, the algorithm terminates in an expected polynomial time in n.

Lemma 3.1.2. *Let $n \geq 8$ be a power of 2 such that $\Phi = x^n + 1$ splits into k_q irreducible factors modulo a prime $q \geq 5$, $\sigma \geq \sqrt{n \log(2n(1 + 1/\delta))/\pi} q^{1/k_q}$, for an arbitrary $\delta \in (0, 1/2)$. Let finally $a \in R$ and $p \in R_q^\times$. Then*

$$\Pr_{f' \sim D_{\mathbb{Z}^n,\sigma}} \left[(pf' + a \mod q) \notin R_q^\times \right] \leq k_q(q^{-n/k_q} + 2\delta) \leq n(q^{-1} + 2\delta).$$

Proof. By the Chinese Remainder Theorem we can proceed by bounding the probability that $pf' + a$ belongs to an ideal $I := \langle q, \Phi_k \rangle$ by $q - n/k_q + 2\delta$, for any $k \leq k_q$. We have $\mathcal{N}(I) = q^{n/k_q}$, so by Minkowski's theorem we get that $\lambda_1(I) \leq \sqrt{n} q^{1/k_q}$. Since I is an ideal of R, we have $\lambda_n(I) = \lambda_1(I)$, and Lemma 1.2.9 gives that $\sigma \geq \eta_\delta(I)$. Finally by Lemma 1.2.14, we obtain that $f \mod I$ is within distance at most 2δ to uniformity on R/I, so we have $pf' + a = 0 \mod I$ with probability at most $q^{-n/k_q} + 2\delta$, as we wanted. To conclude, the union bound given in the previous section leads to the wanted result. □

As a consequence of the bound we just found on the rejection probability, we have also the following result ensuring that the generated secret key is small.

Lemma 3.1.3. *Let $n \geq 8$ be a power of 2 such that $\Phi = x^n + 1$ splits into k_q irreducible factors modulo a prime $q \geq 8n$. Let $\sigma \geq q^{1/k_q}\sqrt{n\mathrm{logn}}$. Then, with probability at least $1 - 2^{-n+3}$, the secret key polynomials f, g returned by the algorithm satisfy the following estimates*

$$\|f\| \leq 2n\|p\|\sigma \quad \textit{and} \quad \|g\| \leq \sqrt{n}\sigma$$

with probability at least $1 - 2^{-n+3}$.

Moreover, if $\deg p \leq 1$, then $\|f\| \leq 4\sqrt{n}\|p\|\sigma$ with probability at least $1 - 2^{-n+3}$.

In the algorithm given above, the polynomials f' and g are independently sampled from the discrete Gaussian distribution $D_{\mathbb{Z}^n,\sigma}$ restricted (by rejection) to $R_q^\times - p^{-1}$ and $R_q^\times$ respectively.

Letting $z \in R_q$, we denote by $D_{\sigma,z}^\times$ the discrete Gaussian $D_{\mathbb{Z}^n,\sigma}$ restricted to $R_q^\times + z$ and $y = -zp^{-1} \mod q$. We want to apply the results of Section 1.5.2 to show the statistical closeness to uniformity of a quotient of two distributions $(z + pD_{\sigma,y}^\times)$, like in the the case of our public key $g/f \mod q$ computed by Algorithm 3.

Since $p \in R_q^\times$, multiplication by p induces an automorphism of R_q, so the statistical closeness to uniformity is preserved on the public key $h = pg/f$. The theorem we are about to see gives two bounds: the first one is most useful for large $k_q = \Omega(n)$, while the second is better for small $k_q = O(1)$, allowing a smaller σ by a factor of the order of $\sqrt{n}$ with respect to the first bound.

Theorem 3.1.4. *Let $n \geq 8$ be a power of 2 such that $\Phi = x^n + 1$ splits into k_q irreducible factors modulo some prime $q \geq 5$, $\varepsilon' \in (0, 1/3)$, $y_i \in R_q$ and $z_i = -y_i p^{-1} \mod q$ for $i = 1, 2$. Then the following two bounds hold:*

(*a*) *if $\sigma \geq n\sqrt{\log(8nq)}q^{\frac{1}{2}+\varepsilon'}$, then*

$$\Delta\left[\frac{y_1 + pD_{\sigma,z_1}^\times}{y_2 + pD_{\sigma,z_2}^\times} \mod q, U(R_q^\times)\right] \leq 2^{10n} q^{-\frac{\lfloor \varepsilon' k_q \rfloor}{k_q} n};$$

(*b*) *if $\sigma \geq \sqrt{n\log(8nq)}q^{\frac{1+k_q\varepsilon'}{2}}$ and $q \geq n^{\frac{k_q}{1-2k_q\varepsilon'}}$, then*

$$\Delta\left[\frac{y_1 + pD_{\sigma,z_1}^\times}{y_2 + pD_{\sigma,z_2}^\times} \mod q, U(R_q^\times)\right] \leq 2^{10n} q^{-\varepsilon' n}.$$

Proof. For $a \in R_q^\times$, define $\Pr_a = \Pr_{f_1,f_2}\left[(y_1 + pf_1)/(y_2 + pf_2) = a\right]$, where $f_i \sim D_{\sigma,z_i}^\times$ for $i = 1, 2$. We want to show that $|\Pr_a - |R_q^\times|^{-1}| \leq \varepsilon''$, where $\varepsilon'' = 2^{2n+5} q^{-n\lfloor \varepsilon' k_q \rfloor / k_q} |R_q^\times|^{-1}$ for (a) and $\varepsilon'' = 2^{6n+4} q^{-\varepsilon' n} |R_q^\times|^{-1}$ for (b). Letting $\mathbf{a} = (a_1, a_2) \in (R_q^\times)^2$, we have that $a_1 f_1 + a_2 f_2 = a_1 z_1 + a_2 z_2$

is equivalent to $(y_1 + pf_1)/(y_2 + pf_2) = -a_2/a_1$ (in $R_q^\times$) and $-a_2/a_1$ is uniformly random in $R_q^\times$ if $\mathbf{a} \sim U((R_q^\times)^2)$. This implies that the fraction of $a \in R_q^\times$ such that $|\Pr_a - |R_q^\times|^{-1}| \leq \varepsilon''$ is equal to the fraction of $\mathbf{a}$ such that $|\Pr_{f_1,f_2}[a_1 f_1 + a_2 f_2 = a_1 z_1 + a_2 z_2] - |R_q^\times|^{-1}| \leq \varepsilon''$.

Moreover, since $(f_1, f_2) = (z_1, z_2) =: \mathbf{z}$ satisfies $a_1 f_1 + a_2 f_2 = a_1 z_1 + a_2 z_2$, the set of solutions $(f_1, f_2) \in R^2$ to the latter equation is $\mathbf{z} + \mathbf{a}^{\perp\times}$, where $\mathbf{a}^{\perp\times} = \mathbf{a}^\perp \cap (R_q^\times + q\mathbb{Z}^n)^2$. This gives

$$\Pr_{f_1,f_2}[a_1 f_1 + a_2 f_2 = a_1 z_1 + a_2 z_2] \\ = \frac{D_{\mathbb{Z}^{2n},\sigma}(\mathbf{z} + \mathbf{a}^{\perp\times})}{D_{\mathbb{Z}^n,\sigma}(z_1 + R_q^\times + q\mathbb{Z}^n) D_{\mathbb{Z}^n,\sigma}(z_2 + R_q^\times + q\mathbb{Z}^n)}.$$

For any $\mathbf{t} \in \mathbf{a}^\perp$, we have $t_2 = -t_1 a_1/a_2$, and since $-a_1/a_2 \in R_q^\times$, t_1 and t_2 must be in the same ideal I_S of R_q for some $S \subseteq \{1, \ldots, k_q\}$. This leads to $\mathbf{a}^{\perp\times} = \mathbf{a}^\perp \setminus \bigcup_{\emptyset \neq S \subseteq \{1,\ldots,n\}} \mathbf{a}^{\perp\times}(I_S)$, and a similar reasoning gives $R_q^\times + \mathbb{Z}^n = \mathbb{Z}^n \setminus \bigcup_{\emptyset \neq S \subseteq \{1,\ldots,n\}} (I_S + q\mathbb{Z}^n)$. By the inclusion-exclusion principle we obtain

$$D_{\mathbb{Z}^{2n},\sigma}(\mathbf{z} + \mathbf{a}^{\perp\times}) = \sum_{S \subseteq \{1,\ldots,n\}} (-1)^{|S|} D_{\mathbb{Z}^{2n},\sigma}(\mathbf{z} + \mathbf{a}^\perp(I_S)), \tag{1}$$

and

$$D_{\mathbb{Z}^n,\sigma}(z_i + R_q^\times + q\mathbb{Z}^n) = \sum_{S \subseteq \{1,\ldots,n\}} (-1)^{|S|} D_{\mathbb{Z}^n,\sigma}(z_i + I_S + q\mathbb{Z}^n). \tag{2}$$

The rest of the proof will be dedicated to show that, except for a fraction at most $2^{9n} q^{-\varepsilon' n}$ of $a \in (R_q^\times)^2$, we have

$$D_{\mathbb{Z}^{2n},\sigma}(\mathbf{z} + \mathbf{a}^{\perp\times}) = (1 + \delta_0)|R_q^\times| q^{-2n} \tag{*}$$
$$D_{\mathbb{Z}^n,\sigma}(z_i + R_q^\times + q\mathbb{Z}^n) = (1 + \delta_i)|R_q^\times| q^{-n} \tag{**}$$

where the $|\delta_i| \leq 2^{2n+2} q^{-n\lfloor \varepsilon' k_q \rfloor / k_q}$ in the first case and $|\delta_i| \leq 2^{6n+1} q^{-\varepsilon' n}$ in the second. Once proved this, the result follows by a straightforward computation.

We will now work separately to get the two bounds just claimed.

Let us start by assuming that (1) holds. First off, since $\mathbf{z} \in \mathbb{Z}^{2n}$, for any $S \subseteq \{1, \ldots, k_q\}$ we have

$$\begin{aligned} D_{\mathbb{Z}^{2n},\sigma}(\mathbf{z} + \mathbf{a}^\perp(I_S)) &= \frac{\rho_\sigma(\mathbf{z} + \mathbf{a}^\perp(I_S))}{\rho_\sigma(\mathbb{Z}^{2n})} \\ &= \frac{\rho_\sigma(\mathbf{z} + \mathbf{a}^\perp(I_S))}{\rho_\sigma(\mathbf{z} + \mathbb{Z}^{2n})} \\ &= D_{\mathbb{Z}^{2n},\sigma,-\mathbf{z}}(\mathbf{a}^\perp(I_S)). \end{aligned}$$

For the terms of (1) with $|S| \leq \varepsilon' k_q$, by setting $\delta = q^{-n(1+\lfloor \varepsilon' k_q \rfloor / k_q)}$ we fall under the assumptions of Lemma 1.5.4. Moreover, since $\mathbf{a} \in (R_q^\times)^2$, there are $q^{n(1-|S|/k_q)}$ elements of $\mathbf{a}^\perp(I_S)$ in $[0, q-1]^{2n}$, and thus $\det(\mathbf{a}^\perp(I_S)) = q^{n(1+|S|/k_q)}$. Using the first bound of Lemma 1.5.4 with $m = 2$ and $\varepsilon = \varepsilon'/2$, we then conclude that

$$\left| D_{\mathbb{Z}^{2n},\sigma,-\mathbf{z}}(\mathbf{a}^\perp(I_S)) - q^{-n(1+|S|/k_q)} \right| \leq 2\delta$$

for all except a fraction at most $2^{8n} q^{-\varepsilon' n}$ of the $\mathbf{a} \in (R_q^\times)^2$.

For a term of (1) with $|S| > \varepsilon' k_q$, we choose $S' \subseteq S$ with $|S'| = \lfloor \varepsilon' k_q \rfloor$. Then we have $\mathbf{a}^\perp(I_S) \subseteq \mathbf{a}^\perp(I_{S'})$ and hence $D_{\mathbb{Z}^{2n},\sigma,-\mathbf{z}}(\mathbf{a}^\perp(I_S) \leq D_{\mathbb{Z}^{2n},\sigma,-\mathbf{z}}(\mathbf{a}^\perp(I'_S))$. By using with S' the above result for small $|S|$, we obtain $D_{\mathbb{Z}^{2n},\sigma,-\mathbf{z}} \mathbf{a}^\perp(I_S) \leq 2\delta + q^{-n(1+\lfloor \varepsilon' k_q \lfloor / k_q)}$. Hence, except possibly for a fraction at most $2^{9n} q^{-\varepsilon' n}$ of $\mathbf{a} \in (R_q^\times)^2$, we have

$$\begin{aligned} \left| D_{\mathbb{Z}^{2n},\sigma}(\mathbf{z} + \mathbf{a}^{\perp\times}) - \sum_{k=0}^{n} (-1)^k \binom{n}{k} q^{-n-k} \right| & \\ \leq 2^{n+1}\delta + 2 \sum_{k=\lceil \varepsilon k_q \rceil} \binom{k_q}{k} q^{-n\left(1+\frac{\lfloor \varepsilon' k_q \rfloor}{k_q}\right)} & \\ \leq 2^{n+1} \left(\delta + q^{-n\left(1+\frac{\lfloor \varepsilon' k_q \rfloor}{k_q}\right)} \right). & \end{aligned}$$

So, in this case, we finally get that

$$|\delta_0| \leq \frac{q^{2n}}{(q^{n/k_q} - 1)^{k_q}} 2^{n+1} (\delta + q^{-n(1+\frac{\lfloor \varepsilon' k_q \rfloor}{k_q})}) \leq 2^{2n+2} q^{-\frac{\lfloor \varepsilon' k_q \lfloor}{k_q} n},$$

which is what we wanted.

Let's now take a look at (b). For the term of (1) with $|S| = 0$, by hypothesis we get that σ falls under the assumptions of Lemma 1.5.4, so we can apply it (in particular: its second part) with $\delta = q^{-2n}$ and $\varepsilon = \varepsilon'/2$. We obtain $|R/\mathbf{a}^\perp(I_S)| = \det(\mathbf{a}^\perp(I_S)) = q^n$ and hence $|D_{\mathbb{Z}^{2n},\sigma,-\mathbf{z}}(\mathbf{a}^\perp(I_S)) - q^{-n}| \leq 2\delta$ for all except a fraction at most $2^{8n} q^{-\varepsilon' n}$ of $\mathbf{a} \in (R_q^\times)^2$.

Assuming $|S| \geq 1$, we need to change our approach. In fact, for $|S| = 1$, we cannot choose an $I_{S'}$ with $S' \subseteq S$ and $\det(\mathbf{a}^\perp(I_{S'}))$ of the order of $q^{(1+\varepsilon)n}$: the only possible choice for S' is the empty set, which gives a too small $\det(\mathbf{a}^\perp(I_{S'})) = q^n$. Let then $L' = N\mathbb{Z}^{2n}$, where $N = \lceil \frac{1}{4} q^{1/2+\varepsilon'/2} \rceil$. Note that $\det L' = N^{2n} \geq 2^{-4n} q^{(1+\varepsilon')n}$, and since $\lambda_{2n}(L') = N \leq \frac{1}{2} q^{1/2+\varepsilon'/2}$, we have by Lemma 1.2.9 with $\delta = q^{-2n}$ that $\eta_\delta(L') \leq \sqrt{n \log(8nq)} q^{1/2+\varepsilon'/2}$. Now, by Lemma 1.2.14 and the choice of σ, we have $D_{\mathbb{Z}^{2n},\sigma}(L') \leq 2^{4n} q^{-(1+\varepsilon')n} + 2\delta$. We want to use this latter bound to conclude; we will do so by showing that $D_{\mathbb{Z}^{2n},\sigma}(\mathbf{z} + \mathbf{a}^\perp(I_S)) \leq D_{\mathbb{Z}^{2n},\sigma}(L')$.

Define the map $\phi : \mathbb{Z}^{2n} \to L'$ via $\mathbf{v} = (v_1, \dots, v_{2n}) \mapsto \phi(\mathbf{v}) = (v'_1, \dots, v'_{2n})$, where $v'_i = \lfloor \frac{|v_i|}{N} \rfloor N \mathrm{sign}(v_i)$. What this function does is namely round each coordinate v_i of $\mathbf{v}$ to the nearest multiple of N whose absolute value is less or equal to $|v_i|$, and it has the following properties:

(i) for each $\mathbf{v} \in \mathbb{Z}^{2n}$, then $\|\phi(\mathbf{v})\| \leq \|\mathbf{v}\|$;

(ii) ϕ is a one-to-one correspondence on $\mathbf{z} + \mathbf{a}^{\perp}(I_S)$ for all but a fraction of at most $2^{4n} q^{-\varepsilon' n}$ of $\mathbf{a} \in (R_q^{\times})^2$.

Whilst (i) follows easily from $|v'_1| \leq |v_1|$ by definition, the property (ii) is less immediate. First off, let's observe that $\|\phi(\mathbf{v}) - \mathbf{v}\|_\infty < N$ for all $\mathbf{v} \in \mathbb{Z}^{2n}$. By contradiction, let us suppose that ϕ is not one-to-one on $\mathbf{z} + \mathbf{a}^{\perp}(I_S)$. Then there exist two vectors $\mathbf{v}_1 \neq \mathbf{v}_2 \in \mathbf{z} + \mathbf{a}^{\perp}(I_S)$ with $\phi(\mathbf{v}_1) = \phi(\mathbf{v}_2)$. By the triangular inequality we have that $\mathbf{v}_1 - \mathbf{v}_2$ is a non-zero vector of $\mathbf{a}^{\perp}(I_S)$ with $\|\mathbf{v}_1 - \mathbf{v}_2\| < 2N \leq q^{1/2+\varepsilon'/2}$. On the other hand, by the first bound of Lemma 1.5.3 with $m = 2$, $|S| = 1$, and $\varepsilon = \varepsilon'/2$, we have $\lambda_1^{\infty}(\mathbf{a}^{\perp}(I_S)) \leq \frac{1}{\sqrt{n}} q^{\frac{1}{2} + \frac{1}{2k_q} - \frac{\varepsilon'}{2}}$, except for a fraction at most $2^{4n} q^{-\varepsilon' n}$ of $\mathbf{a} \in (R_q^{\times})^2$. This contradicts the condition on q, thus giving us (i).

Since $D_{\mathbb{Z}^{2n},\sigma}(\mathbf{w}) \geq D_{\mathbb{Z}^{2n},\sigma}(\mathbf{v})$ for any $\mathbf{v}, \mathbf{w} \in \mathbb{Z}^{2n}$ with $\|\mathbf{w}\| \leq \|\mathbf{v}\|$, the property (i) of ϕ implies that $D_{\mathbb{Z}^{2n},\sigma}(\mathbf{z} + \mathbf{a}^{\perp}(I_S)) \leq \sum_{\mathbf{v} \in \mathbf{z} + \mathbf{a}^{\perp}(I_S)} (\phi(\mathbf{v}))$, and by the property (ii) we get that the points $\{\phi(\mathbf{v})\}_{\{\mathbf{v} \in \mathbf{z} + \mathbf{a}^{\perp}(I_S)\}}$ are distinct points of L', so that $\sum_{\mathbf{v} \in \mathbf{z} + \mathbf{a}^{\perp}(I_S)} (\phi(\mathbf{v})) \leq D_{\mathbb{Z}^{2n},\sigma}(L')$, as required.

To conclude, for the terms with $|S| \geq 1$, we have $D_{\mathbb{Z}^{2n},\sigma,-\mathbf{z}}(\mathbf{a}^{\perp}(I_S)) \leq 2^{4n+1} q^{-(1+\varepsilon')n}$. Arguing analogously to what we did to get the first bound, we obtain our second bound $|\delta_0| \leq \frac{q^{2n}}{(q^{n/k_q} - 1)^{k_q}} 2^{5n+1} q^{-(1+\varepsilon)')n} \leq 2^{6n+1} q^{-\varepsilon' n}$.

Let's take care of (2). For the bounds on δ_1 and δ_2, we want to proceed in a similar way to handle the z_i's, *i.e.* we look for a bound on $D_{\mathbb{Z}^n,\sigma,-z_i}(I_S + q\mathbb{Z}^n)$, and by Lemma 1.2.14 this reduces to finding a good bound on the smoothing parameter of the ideal lattice $L_S = I_S + q\mathbb{Z}^n$.

First, observe that if $m = 1$ and $a_1(x) = \prod_{i \in \overline{S}} \Phi_i(x)$, with $\overline{S} = \{1, \dots, n\} \setminus S$, we have $L_S = \mathbf{a}^{\perp}(I_S)$. Then, since $a_1(x) \mapsto a_1^{\star}(x)$ induces a bijection on the factors $\Phi_i(x)$, by Lemma 1.5.1 the dual lattice $L_S^{\vee} = \frac{1}{q} L(a_1^{\star}, I_{\overline{S}}^{\star}) = \frac{1}{q} L_{\overline{S}'}^{\star}$ is also an ideal lattice for some $\overline{S}' \subseteq \{1, \dots, k_q\}$, with $|\overline{S}'| = |\overline{S}|$. Now, since $\det L_{\overline{S'}} = q^{n|\overline{S}|/k_q}$, Minkowski's theorem gives that $\lambda_1^{\infty}(L_{\overline{S}'}) \leq q^{|\overline{S}|/k_q}$. Moreover, since $I_S + q\mathbb{Z}^n$ is an ideal lattice, for $\delta = q^{-n/2}$ and $|S| \leq k_q/2$, Lemma 1.2.9 gives that

$$\eta_\delta(I_S + q\mathbb{Z}^n) \leq \frac{1}{q} \sqrt{\log(2n(1 + 1/\delta))/\pi} \lambda_1^{\infty}(L_{\overline{S}'}) \leq \sqrt{n \log(4nq)} q^{|S|/k_q} \leq \sigma.$$

Using Lemma 2.4, we conclude that for a term of (2) with $|S| \leq k_q/2$, we have

$$\left| D_{\mathbb{Z}^n,\sigma,-z_i}(I_S + q\mathbb{Z}^n) - q^{-n|S|/k_q} \right| \leq 2\delta.$$

For terms of (2) with $|S| > k_q/2$, we choose $S' \subseteq S$ with $|S'| = \lfloor k_q/2 \rfloor \geq k_q/3$ for $k_q \geq 2$. Using the above result for small $|S|$ with S', we obtain $D_{\mathbb{Z}^n,\sigma,-z_i}(I_S + q\mathbb{Z}^n) \leq D_{\mathbb{Z}^n,\sigma,-z_i}(I_{\overline{S}'} + q\mathbb{Z}^n) \leq 2\delta + q^{-n/3}$. Summarizing, we have

$$\begin{aligned} \left| D_{\mathbb{Z}^n,\sigma}(z_i + R_q^\times + q\mathbb{Z}^n) - \sum_{k=0}^{k_q} (-1)^k \binom{k_q}{k} q^{-k} \right| \\ \leq 2^{n+1}\delta + 2 \sum_{k=\lceil k_q/2 \rceil}^{k_q} \binom{k_q}{k} q^{-n/3} \\ \leq 2^{n+1}\big(\delta + q^{-n/3}\big), \end{aligned}$$

which leads to the desired bound on δ_i. □

3.2 A revised NTRUEncrypt scheme

In this section we use the new key generation algorithm to build a provably secure variant of the NTRUEncrypt scheme. We will denote the new scheme with NTRUEncrypt$(n, q, p, \alpha, \sigma)$ in which the parameters are chosen in the following way:

- n is a power of 2;
- $q > 3$ is a prime;
- define $\Phi = x^n + 1$, $R = \mathbb{Z}[x]/\Phi$, $R_q = R/qR$;
- $p \in R_q^\times$, from which we get the plaintext message space as $\mathcal{P} = R/pR$;
- α is the R-LWE noise distribution parameter;
- σ is the standard deviation of the discrete Gaussian distribution used in the key generation process defined in the previous section.

There is to note that p must be a polynomial with small coefficients with respect to q, but at the same time we require $\mathcal{N}(p) = |\mathcal{P}| = 2^{\Omega(n)}$ to be able to encode multiple bits at once. By reducing modulo the px^i's, any element of $\mathcal{P}$ can be written as $\sum_{0 \leq i < n} \varepsilon_i x^i p$, with $\varepsilon_i \in (-1/2, 1/2]$. We can then assume that any element of $\mathcal{P}$ is an element of R with infinity norm at most $\frac{1}{2}\sqrt{\deg(p) + 1}\|p\|$.

We are ready to give the new NTRUEncrypt scheme:

Algorithm 4 Encryption

```
Input: The parameters $n, q, p, \alpha, \sigma$, a message $M \in \mathcal{P}$
Sample $s, e$ from $\overline{\Upsilon}_\alpha$
return $C = hs + pe + M \in R_q$
Output: The ciphertext $C$
```

For the decryption process, it is enough to compute $C' = fC \in R_q$ and return $M = C' \mod p$.

Let us now prove the correctness of this scheme and the actual conditions for it to be sound.

Lemma 3.2.1. *If* $\deg p \leq 1$, $\omega(n^{\frac{1}{4}} \log n)\alpha\|p\|^2\sigma < 1$, *and* $\alpha q \geq n^{\frac{3}{4}}$, *then the decryption algorithm of NTRUEncrypt recovers* M *with probability* $1 - n^{-\omega(1)}$ *over the choice of* s, e, f, g.

Proof. Let $C'' = p(gs + ef) + fM$ in R, *i.e.* before the reduction modulo q that would give C' as in the decryption step. If $\|C''\|_\infty < q/2$, then $C' = C''$ in R. This means that, since $f = 1 \mod p$, we would have $C' \mod p = C'' \mod p = M \mod p$, that is a success in the decryption. For this reason it will be enough to give an upper bound on the probability that $\|C''\|_\infty > q/2$.

For $\deg p \leq 1$, we know from Lemma 3.1.3 that both f and g have Euclidean norms at most $4\sqrt{n}\|p\|\sigma$ with probability not less than $1 - 2^{-n+3}$, so $\|pf\|, \|pg\| \leq 8\sqrt{n}\|p\|^2\sigma$, with probability at least $1 - 2^{-n+3}$. Now, we know that $\|fM\|_\infty \leq \|fM\| \leq \sqrt{n}\|f\|\|M\| \leq 4n\|p\|^2\sigma$, and moreover Lemma 2.2.17 grants that both $\|pfs\|_\infty, \|pge\|_\infty \leq 8\alpha q n^{\frac{1}{4}}\omega(\log n)\|p\|^2\sigma$ with probability at least $1 - n^{-\omega(1)}$.

From these two facts, since $\alpha q \geq n^{\frac{3}{4}}$, we conclude that

$$\|C''\|_\infty \leq 20\alpha q n^{\frac{1}{4}}\omega(\log n)\|p\|^2\sigma$$

with probability $1 - n^{-\omega(1)}$. □

With Theorem 3.1.4 we proved that the public key is basically uniform in $R_q^\times$. Now we want to use such result to reduce the security of our scheme to a decisional instance of R-LWE$^\times_{\text{HNF}}$.

Lemma 3.2.2. *Let* n *be a power of* 2 *such that* Φ *splits into* n *linear factors modulo* q, $\varepsilon \in (0, 1/3)$, $\delta > 0$, $p \in R_q^\times$ *and* $\sigma \geq n\sqrt{\log(8nq)}q^{\frac{1}{2}+\varepsilon}$. *If there exists an IND-CPA attack against NTRUEncrypt that runs in time* T *with success probability* $\frac{1}{2} + \delta$, *then there exists an algorithm that solves R-LWE*$^\times_{HNF}$ *with parameters* q *and* α *running in time* $T' = T + O(n)$ *and with success probability* $\delta' = \delta - q^{-\Omega(n)}$.

Proof. Let's call $\mathcal{A}$ the IND-CPA attack algorithm given in the statement; we will use it to build an algorithm $\mathcal{B}$ against R-LWE$\times_{\text{HNF}}$. Let $\psi \sim \overline{\Upsilon}_\alpha$, $s \sim \psi$ and O be an oracle that samples from either $U(R_q^\times \times R_q)$ or $A_{s,\psi}^\times$.

First, algorithm $\mathcal{B}$ would call the oracle O to get a sample $(h', C') \in R_q^\times \times R_q$ and use $h = ph' \in R_q$ as a public key to run algorithm $\mathcal{A}$. This would output challenge messages $M_0, M_1 \in \mathcal{P}$, and algorithm $\mathcal{B}$ picks $b \sim U(\{0,1\})$, computes the challenge ciphertext $C = pC' + M_b \in R_q$, and returns C to $\mathcal{A}$. So in the end, $\mathcal{A}$ outputs its guess b' for b, and $\mathcal{B}$ outputs 1 if $b' = b$ and 0 otherwise.

We should observe that since p is invertible modulo q, the h' used by $\mathcal{B}$ is uniformly random in $R_q^\times$, and so is the public key h given to $\mathcal{A}$. This implies (by Theorem 3.1.4) that the public key given to $\mathcal{A}$ is within statistical distance $q^{-\Omega(n)}$ from the public key distribution in the attack. Moreover, the ciphertext C given to $\mathcal{A}$ has the right distribution as in the IND-CPA attack, because $C' = hs + e$ with $s, e \sim \psi$. So if the output of O is a sample from $A_{s,\psi}^\times$, then $\mathcal{A}$ succeeds and $\mathcal{B}$ returns 1 with probability at least $1/2 + \delta - q^{-\Omega(n)}$.

On the other hand, if O outputs samples from $U(R_q^\times \times R_q)$, the value of pC' – and hence C – is uniformly random in R_q and independent of b, because $p \in R_q^\times$. This means algorithm $\mathcal{B}$ outputs 1 with probability $1/2$.

Putting together the two possibilities we get the advantage we claimed for $\mathcal{B}$. □

We finally get for free to our main result:

Theorem 3.2.3. *Let n be a power of 2 for which Φ splits into n linear factors modulo a prime $q = poly(n)$ such that $q^{\frac{1}{2}-\varepsilon} = \omega(n^{\frac{9}{8}} \log 2n)\|p\|^2$, with $\varepsilon = \omega(1/n) < 1/3$ and $p \in R_q^\times$ with $\deg(p) \leq 1$. Let moreover $\sigma = n\sqrt{\log(8nq)}q^{\frac{1}{2}+\varepsilon}$ and $\alpha^{-1} = \omega(n^{\frac{1}{4}} \log n)\|p\|^2\sigma$. If there exists an IND-CPA attack against NTRUEncrypt running in polynomial time in n and with success probability at least $\frac{1}{2} + 1/poly(n)$, then there exists a $poly(n)$-time quantum algorithm for Ideal-SVP$_\gamma$ with $\gamma = \omega(n^{\frac{11}{8}} \log^{\frac{5}{2}} n)\|p\|^2 q^{\frac{1}{2}+\varepsilon}$. Moreover, the decryption algorithm succeeds with probability $1 - n^{-\omega(1)}$.*

Proof. The result follows directly by the Lemmata above and Theorem 2.2.15. □

For a choice of $\varepsilon = 1/(\log n)$, the smallest q for which this analysis holds is of the order of $\widetilde{\Omega}(n^{\frac{9}{2}})$, and the smallest value of γ we can obtain is $\widetilde{O}(n^5)$. Finally, the problem of finding the best set of parameters to make this encryption practical while still satisfy the security hypothesis is still open.

Bibliography

[Ban93] Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(1):625–635, 1993.

[CHKP12] David Cash, Dennis Hofheinz, Eike Kiltz, and Chris Peikert. Bonsai trees, or how to delegate a lattice basis. *J. Cryptol.*, 25(4):601–639, October 2012.

[GGH97] Oded Goldreich, Shafi Goldwasser, and Shai Halevi. Public-key cryptosystems from lattice reduction problems. In *Advances in Cryptology-CRYPTO'97: 17th Annual International Cryptology Conference, Santa Barbara, California, USA, August 1997. Proceedings*, page 112. Springer, 1997.

[GPV08] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 197–206. ACM, 2008.

[HG07] Nick Howgrave-Graham. A hybrid lattice-reduction and meet-in-the-middle attack against ntru. In *Proceedings of the 27th Annual International Cryptology Conference on Advances in Cryptology*, CRYPTO'07, pages 150–169. Springer, 2007.

[HPS98] Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman. NTRU: A ring-based public key cryptosystem. In *Lecture Notes in Computer Science*, pages 267–288. Springer, 1998.

[HPSS08] Jeffrey Hoffstein, Jill Pipher, Joseph H. Silverman, and Joseph H. Silverman. *An introduction to mathematical cryptography*. Springer, 2008.

[LM09] Vadim Lyubashevsky and Daniele Micciancio. On bounded distance decoding, unique shortest vectors, and the minimum distance problem. In *Halevi S. (eds) Advances in Cryptology - CRYPTO 2009, Lecture notes in Computer Science*, volume 5677, pages 577–594. Springer, 2009.

[LPR12] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 1–23. Springer, 2012.

[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.

[McE78] Robert J. McEliece. A public-key cryptosystem based on algebraic coding theory. *The Deep Space Network Progress Report, January and February 1978*, 42-44:114–116, 1978.

[Mic07] Daniele Micciancio. Generalized compact knapsacks, cyclic lattices, and efficient one-way functions. *Computational Complexity*, 16(4):365–411, 2007.

[MR08] Daniele Micciancio and Oded Regev. Lattice-based cryptography. In Daniel J. Bernstein, Johannes Buchmann and Erik Dahmen (eds)*, Post-Quantum Cryptography*, pages 147–191, 2008.

[Pei09] Chris Peikert. Public-key cryptosystems from the worst-case shortest vector problem. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 333–342. ACM, 2009.

[Pei16] Chris Peikert. A decade of lattice cryptography. *Foundations and Trends in Theoretical Computer Science*, 10(4):283–424, 2016.

[PW11] Chris Peikert and Brent Waters. Lossy trapdoor functions and their applications. *SIAM Journal on Computing*, 40(6):1803–1844, 2011.

[Reg09] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *Journal of the ACM*, 56(6):34, 2009.

[SS11] Damien Stehlé and Ron Steinfeld. Making NTRUEncrypt and NTRUSign as secure as standard worst-case problems over ideal lattices. Paterson K.G. (eds) Advances in Cryptology – EUROCRYPT 2011, 2011. Updated version (2013): IACR Cryptology ePrint Archive, Report 2013/004 `http://eprint.iacr.org/2013/004`.