Revisiting Lattice Attacks on overstretched NTRU parameters

Paul Kirchner[1,2] and Pierre-Alain Fouque[2,3]

[1] École normale supérieure
[2] IRISA
[3] Université de Rennes 1 & Institut Universitaire de France
pkirchner@clipper.ens.fr, pierre-alain.fouque@univ-rennes1.fr

Abstract. In 2016, Albrecht, Bai and Ducas and independently Cheon, Jeong and Lee presented very similar attacks to break the NTRU cryptosystem with larger modulus than in the `NTRUEncrypt` standard. They allow to recover the secret key given the public key of Fully Homomorphic Encryption schemes based on NTRU ideas. Hopefully, these attacks do not endanger the security of the `NTRUEncrypt`, but shed new light on the hardness of the NTRU problem. The idea consists in decreasing the dimension of the NTRU lattice using the multiplication matrix by the norm (resp. trace) of the public key in some subfield instead of the public key itself. Since the dimension of the subfield is smaller, so is the dimension of the lattice and better lattice reduction algorithms perform. In this paper, we first propose a new variant of the subfield attacks that outperforms both of these attacks in practice. It allows to break several concrete instances of YASHE, a NTRU-based FHE scheme, but it is not as efficient as the hybrid method on smaller concrete parameters of `NTRUEncrypt`. Instead of using the norm and trace, the multiplication by the public key in a subring allows to break smaller parameters and we show that in $\mathbb{Q}(\zeta_{2^n})$, the time complexity is polynomial for $q = 2^{\Omega(\sqrt{n \log\log n})}$. Then, we revisit the lattice reduction part of the hybrid attack of Howgrave-Graham and analyze the success probability of this attack using a new technical tool proposed by Pataki and Tural. We show that, under some heuristics, this attack is more efficient than the subfield attack and works in any ring for large q, such as the NTRU Prime ring. We insist that the improvement on the analysis applies even for relatively small modulus ; although if the secret is sparse, it may not be the fastest attack. We also derive a tight estimation of security for (Ring-)LWE and NTRU assumptions and perform many practical experiments.

1 Introduction

NTRU has been introduced by Hoffstein, Pipher and Silverman since 1996 in [26] and has since resisted many attacks [13,22,21,27]. NTRU is one of the most attractive lattice-based cryptosystems since it is very efficient, and many Ring-LWE cryptosystems have a NTRU variant. Ducas, Lyubashevsky and Prest propose an Identity Based Encryption scheme based on NTRU [20] (albeit with a

much larger standard deviation), López-Alt, Tromer and Vaikuntanathan describe a Fully Homomorphic Encryption scheme [32], which is improved in a scheme called YASHE [6,31], and Ducas *et al.* propose a very fast signature scheme called BLISS [19].

Currently, the most efficient and heuristic attack on NTRU has been given by Kirchner and Fouque in [29] which has subexponential-time complexity in $2^{(n/2+o(n))/\log\log q}$, but the $o(n)$ is too large to lead to attack for given parameters. To date, the most efficient attack on practical NTRU parameters is the so-called hybrid attack described by Howgrave-Graham in [27].

The key recovery problem of NTRU is the following problem: given a public key $\mathbf{h} = \mathbf{f}/\mathbf{g}$ in some polynomial ring $R_q = \mathbb{Z}_q[X]/(X^n - 1)$ for n prime, q a small integer and the euclidean norms of $\mathbf{f}, \mathbf{g}$ are small, recover $\mathbf{f}$ and $\mathbf{g}$ or a small multiple of them. In `NTRUEncrypt`, $\mathbf{f}$ and $\mathbf{g}$ are two sparse polynomials of degrees strictly smaller than n and coefficients $\{-1, 0, 1\}$. It is easy to see that the public key cannot be uniformly distributed in the whole ring, since the entropy is too small. In [43], Stehlé and Steinfeld, show that if $\mathbf{f}$ and $\mathbf{g}$ are generated using a Gaussian distribution of standard deviation $\sigma \approx q^{1/2}$, then the distribution of the public key is statistically indistinguishable from the uniform distribution.

State-of-the-art lattice algorithm on NTRU. In [13], Coppersmith and Shamir show that the $(2n)$-dimensional lattice L^{cs} generated by the columns of the matrix

$$\begin{pmatrix} q\mathbf{I}_n & \mathbf{M}_{\mathbf{h}}^{R_q} \\ \mathbf{0} & \mathbf{I}_n \end{pmatrix},$$

where $\mathbf{M}_{\mathbf{h}}^{R_q}$ denotes the multiplication by the public key $\mathbf{h}$ in the ring R_q, contains the vector $(\mathbf{f}, \bar{\mathbf{g}})$. It is easy to show that for $\bar{\mathbf{g}} = \mathbf{g}(1/X)$ in R_q, we have $\mathbf{h} \cdot \bar{\mathbf{g}} = \mathbf{f}$. By reducing this lattice, it is possible to find $(\mathbf{f}, \bar{\mathbf{g}})$ which is short if $(\mathbf{f}, \mathbf{g})$ is. Finally, Coppersmith and Shamir show that for cryptographic purposes, it is sufficient to recover a small solution, maybe not the smallest one to decrypt.

In 2001, May showed in [33] how to exploit that the shifts of the target vector, i.e. $\mathbf{x}^i \cdot \mathbf{f}$ in R_q are also contained in the L^{cs} lattice. Consequently, we only have to recover one of the n shifts of the target vector and the smallest vector is not unique. The idea of May consists in constructing a lattice that contains as a short vector only one of the shift and such that the gap between the first and second minima of the lattice will be higher. This gap is an important parameter when running lattice reduction algorithm. If we take into account that the vector of the secret key contains $\{0, \pm 1\}$-coefficients, there is a unique long run of 0-coefficients. For one of the n shifts, we can assume that this run is for instance in the first r coefficients and if we multiply the $(n+1)$th to $(n+r)$th columns of L^{cs} matrix by a suitable large constant, only this shift will be a solution for the new lattice. He also introduces the *projection* technique to reduce the dimension of L^{cs} from $2n$ to $(1+\alpha)n$ for $0 < \alpha \leq 1$ by removing the last columns of the matrix $\mathbf{M}_{\mathbf{h}}^{\mathcal{O}_{\mathbb{K}}}$ or of the last rows of the original matrix. The main idea is that it suffices that among the n equations corresponding to $\mathbf{h} \cdot \bar{\mathbf{g}} = \mathbf{f}$, some of them will not be fulfilled. Experimentally, since there is no other small vector except the n

shifts, then there will be no other vector with small entries in these coefficients and we will recover the target.

In [27], Howgrave-Graham makes various experiments on NTRU lattice and proposes a mix between lattice reduction and a combinatorial technique, known as Odlyzko's meet-in-the-middle attack on NTRU. The first phase of the algorithm starts by reducing the original matrix corresponding to L^{cs} and we can see that lattice algorithms first reduce the column vectors in the middle of the matrix. This process that treats the columns in a symmetric manner between $[n-r, n+r]$ is also used in [21] in the symplectic reduction. Consequently, it is more efficient to begin by reducing a small dimensional matrix in the center of the original Coppersmith-Shamir matrix and then another combinatorial technique can take into account the small coefficients in the short vector by guessing some part of the secret key. In the following, we will speak of the *middle* technique.

More recently, in [12,1], Cheon, Jeong and Lee at ANTS 2016 and Albrecht, Bai and Ducas at CRYPTO 2016, described a new attack on NTRU-like cryptosystems. An attack based on similar ideas was proposed by Jonsson, Nguyen and Stern in [23, Section 6]. It uses the fact that for cyclotomic number fields, there exist subfields that allow to reduce the dimension of the lattice. The *subfield* attack recovers the norm of the secret key in these subfields, which are smaller than in the classical NTRU lattice. In the maximal real subfield $\mathbb{K}^+$ of a power of two cyclotomic field $\mathbb{K}$ for instance, the norm can be written as $\mathsf{N}_{\mathbb{K}/\mathbb{K}^+}(\mathbf{f}) = \mathbf{f}\bar{\mathbf{f}}$ which is small if $\mathbf{f}$ is small and $\mathsf{N}_{\mathbb{K}/\mathbb{K}^+}(\mathbf{f})$ is of dimension half. The lattice L^{norm} is generated by the columns of the matrix of dimension n:

$$\begin{pmatrix} q\mathbf{I}_{n/2} & \mathbf{M}_{\mathbf{h}\bar{\mathbf{h}}}^{\mathcal{O}_{\mathbb{K}^+}} \\ \mathbf{0} & \mathbf{I}_{n/2} \end{pmatrix}.$$

The vector $(\mathsf{N}_{\mathbb{K}/\mathbb{K}^+}(\mathbf{f}), \mathsf{N}_{\mathbb{K}/\mathbb{K}^+}(\mathbf{g}))$ is small in L^{norm}. By the Gaussian heuristic, the expected length of the shortest vector in the lattice L^{norm} is $\sqrt{qn/(2\pi e)}$, and the norm of $\mathbf{f}$ depends on the density of non-zero coefficients is of size around n. For standard NTRU parameters and when n is greater than q, lattice reduction algorithms will not recover the secret key. However, if q is large as in the case of FHE cryptosystems to allow a large number of multiplication steps before boostraping, then this attack can be interesting. We have not been able to apply it for other cryptosystems, for instance on IBE and signature schemes [20].

The drawback of this technique is that q has to be very large compared to n. We estimate asymptotically $q = 2^{\Omega(\sqrt{n \log\log n})}$ for a polynomial time complexity.

Our Results. In this paper, we revisit the lattice attacks on NTRU by considering the subfield idea, the projection of May and the middle lattice of Howgrave-Graham in the context of large modulus.

1. We first propose a new *subfield* attack and give, contrary to [1,12], a precise analysis by considering the *projection* technique for power of two cyclotomic fields. We show that using the multiplication matrix by the public key in

a subring (which has the same size as the subfield), leads to more efficient attacks. In particular, we were able to attack concrete parameters proposed in YASHE based on overstretched NTRU [6,7,30,15,16,14], [10,31], meaning that we can recover a decryption key for smaller modulus q, compared to the previous approaches [1,12]. The previous attacks use the norm over the subfield in [1] or the trace in [12]. It would also be possible for instance to use all the coefficients of the characteristic polynomial. Our attack using the subring is better than the two previous ones since in the same configuration, we can choose exactly the size of the subfield as the number of coordinates (remove some rows or *project* the lattice) in Section 3.

2. Secondly, we analysis lattice reduction algorithm on the full L^{cs} lattice using a nice lemma due to Petaki and Tural [39] on the volume of sublattices with high rank (Section 4). We show that reducing this lattice allows us to achieve similar performances as in the projection and subfield attacks. We do not rely in our analysis on the Hermite factor (or approximate factor). This is the first time that the high number of small lattice vectors (shifts) are used to improve the analysis of the attack against NTRU. May used it to run lattice reduction on smaller dimensional lattices. The high dimensional low volume sublattice (formed by the shift vectors) makes the approximate-SVP problem for NTRU lattices substantially easier to solve by lattice reduction than generic lattices of the same dimension when the modulus is sufficiently large. This result is true in any ring and can be applied for instance on NTRUPrime with large q. In practice, we run experiment using the *middle* technique in order to use small dimension lattices.

3. We make experiments (Section 5) to understand the behaviour of lattice reduction algorithm and derive precise predictions when this attack will work (Section 6). We show that also experimentally the subfield attack is not more efficient than the *middle* technique on the original matrix. Consequently, we mount this attack to break FHE with NTRU and overstretched NTRU Prime scheme. Experimental computations show that if we are able to reduce this matrix, we recover a basis consisting of n small vectors, which are rotated version of the secret key. Finally, we provide a tight asymptotical security estimate of NTRU and LWE schemes in order to give exact predictions for these attacks by considering the Dual-BKZ [37].

We want to stress that the subfield attack we propose is not needed to break the schemes. We first discovered our subfield attack and the experiments shown in Figure 1 have been obtained using it. The experiments on NTRUPrime with overstretched parameters (Figure 2) have been achieved by reducing the middle lattice in the standard lattice. We experimentally recovered the same results for Figure 1 using the middle lattice later and we conclude that the subfield attack is not needed to improve results on NTRU, but it could be useful to attack multilinear maps [12,1].

2 Preliminaries

Algebraic number field. An *algebraic number field* (or simply number field) K is a finite (algebraic) field extension of the field of rational numbers $\mathbb{Q}$. An *algebraic number* $\zeta \in \mathbb{C}$ is a root of a polynomial $f(x) \in \mathbb{Q}[x]$ and is called an *algebraic integer* if $f(x)$ is a monic (leading coefficient is 1), polynomial in $\mathbb{Z}[x]$. The *minimal polynomial* of ζ is a monic polynomial $f(x) \in \mathbb{Q}[x]$ of least positive degree such that $f(\zeta) = 0$ and the minimal polynomial of an algebraic integer is in $\mathbb{Z}[x]$. The set of all algebraic integers form a ring: the sum and product of two algebraic integers is an algebraic integer. The *ring of integers of a number field* $K = \mathbb{Q}[\zeta]$, obtained by adjoining ζ to $\mathbb{Q}$, is the ring $\mathcal{O}_K = \{\mathbf{x} \in K : \mathbf{x} \text{ is an algebraic integer}\}$. Let $f(x)$ be the minimal polynomial of ζ of degree n, then as $f(\zeta) = 0$, there is an isomorphism between $\mathbb{Q}[x]$ mod $f(x)$ and K, defined by $x \mapsto \zeta$ and K can be seen as an n-dimensional vector space over $\mathbb{Q}$ with *power basis* $\{1, \zeta, \ldots, \zeta^{n-1}\}$. The *conjugates* of ζ are defined as all the roots of its minimal polynomial.

A number field $K = \mathbb{Q}[\zeta]$ of degree n has exactly n field homomorphisms $\sigma_i : K \hookrightarrow \mathbb{C}$ that fix every element of $\mathbb{Q}$ and they map ζ to each of its conjugates. An embedding whose image lies in $\mathbb{R}$ (real root of $f(x)$) is called a *real embedding*; otherwise it is called a *complex embedding*. Since complex root of $f(x)$ come in pairs, so do complex embeddings. The number of real ones is denoted s_1 and the number of pairs of complex ones s_2, so we get $n = s_1 + 2s_2$. By convention, we let $\{\sigma_j\}_{j\in[s_1]}$ be the real embedding and order the complex embeddings so that $\sigma_{s_1+s_2+j} = \overline{\sigma_{s_1+j}}$ for $j \in [s_2]$. The *canonical embedding* $\sigma : K \to \mathbb{R}^{s_1} \times \mathbb{C}^{2s_2}$ is defined by

$$\sigma(\mathbf{x}) = (\sigma_1(\mathbf{x}), \ldots, \sigma_n(\mathbf{x})).$$

The canonical embedding σ is a field homomorphism from K to $\mathbb{R}^{s_1} \times \mathbb{C}^{2s_2}$, where multiplication and addition in $\mathbb{R}^{s_1} \times \mathbb{C}^{2s_2}$ are component-wise. The discriminant Δ_K of K is the determinant of the matrix $(\sigma_i(\alpha_j))_{i,j}$, where (α_j) is a set of n elements of K.

For elements $H \subseteq \mathbb{R}^{s_1} \times \mathbb{C}^{2s_2} \subset \mathbb{C}^n$ where

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

we can identify elements of K to their canonical embeddings in H and speak of the geometric canonical norms on K as $\|\mathbf{x}\|$ as $\|\sigma(\mathbf{x})\|_2 = (\sum_{i\in[n]} |\sigma_i(\mathbf{x})|^2)^{1/2}$. The field norm of an element $\mathbf{a} \in K$ is defined as $\mathsf{N}_{K/\mathbb{Q}}(\mathbf{a}) = \prod_{i\in[n]} \sigma_i(\mathbf{a})$. Note that the norm of an algebraic integer is in $\mathbb{Z}$ as the constant coefficient of the minimal polynomial. Let L a subfield of K, the relative norm of $\mathsf{N}_{K/L}(\mathbf{a}) = \prod_{\sigma_i\in\mathrm{Gal}(K/L)} \sigma_i(\mathbf{a})$, where $\mathrm{Gal}(K/L)$ contains the elements that fix L. The trace of $\mathbf{a} \in K$ is defined $\mathsf{Tr}_{K/\mathbb{Q}}(\mathbf{a}) = \sum_{i\in[n]} \sigma_i(\mathbf{a})$ and is the trace of the endomorphism $\mathbf{y} \mapsto \mathbf{ay}$ and of its matrix representation.

Let K a number field of dimension n, which has a subfield L of dimension $m \mid n$. For simplicity, we assume that K is a Galois extension of $\mathbb{Q}$, with Galois group G ; and G' is the subgroup of G fixing L. It is a standard fact that $|G'| = n/m$.

Notice that elements of the Galois group permute or conjugate the coordinates in $\mathbb{R}^r \times \mathbb{C}^s$, and therefore the norm is invariant by elements of G:

$$\forall \sigma \in G, \|\sigma(\mathbf{x})\| = \|\mathbf{x}\|.$$

We call $\mathsf{N}_{K/L} : K \to L$ the relative norm, with $\mathsf{N}_{K/L}(\mathbf{a})$ the determinant of the L-linear endomorphism $\mathbf{x} \mapsto \mathbf{a}\mathbf{x}$. It is known that we have:

$$\mathsf{N}_{K/L}(\mathbf{a}) = \prod_{\sigma \in H} \sigma(\mathbf{a}).$$

We can bound the norm using the inegality of arithmetic and geometric means:

$$|\mathsf{N}_{K/\mathbb{Q}}(\mathbf{a})| \leq \left(\frac{\|\mathbf{a}\|}{\sqrt{n}}\right)^n.$$

The operator norm for the euclidean norm is denoted $\|\cdot\|_{op}$ and is defined as $\|\mathbf{a}\|_{op} = \sup_{\mathbf{x} \in \mathbb{K}^*} \|\mathbf{a}\mathbf{x}\|/\|\mathbf{x}\|$. Remark that it is simply the maximum of the norm of the coordinates in $\mathbb{R}^r \times \mathbb{C}^s$. Also, it is sub-multiplicative and $\|\mathbf{x}\| \leq \sqrt{n}\|\mathbf{x}\|_{op}$.

Let $\mathcal{O}$ be an *order* of K, that is $\mathcal{O} \subset K$ and $\mathcal{O}$ is a commutative group which is isomorphic as an abelian group to $\mathbb{Z}^n$. We define $\mathcal{O}_L$ as $\mathcal{O} \cap L$, and is an order of L. We denote by $\mathrm{Vol}(\mathcal{L})$ the volume of the lattice $\mathcal{L}$, which is the square root of the determinant of the Gram matrix corresponding to any basis of $\mathcal{L}$. We define Δ to be the square of the volume of $\mathcal{O}$, and likewise for Δ_L with respect to $\mathcal{O}_L$.

We define

$$\mathbf{M}_{\mathbf{a}}^{\mathcal{L}} : \begin{array}{c} \mathcal{L} \longrightarrow \mathcal{O} \\ \mathbf{x} \longmapsto \mathbf{a}\mathbf{x} \end{array}$$

for any lattice $\mathcal{L} \subset \mathcal{O}$ and $\mathbf{a} \in \mathcal{O}$; and we also denote $\mathbf{M}_{\mathbf{a}}^{\mathcal{L}}$ the corresponding matrix for some basis of $\mathcal{L}$.

Cyclotomic field. In the case of cyclotomic field defined by $\Phi_{\mathfrak{f}}(x) = \prod_{k \in \mathbb{Z}_{\mathfrak{f}}^*}(x - \zeta_{\mathfrak{f}}^k)$, where $\zeta_{\mathfrak{f}} = e^{2i\pi/\mathfrak{f}} \in \mathbb{C}$, a primitive $\mathfrak{f}$-root of unity. Thus, $\Phi_{\mathfrak{f}}(x)$ has degree $n = \varphi(\mathfrak{f})$, is monic and irreducible over $\mathbb{Q}$ and its the minimal polynomial of the algebraic integer $\zeta_{\mathfrak{f}}$. The $\mathfrak{f}$th *cyclotomic field* is $\mathbb{Q}[\zeta_{\mathfrak{f}}]$ and its ring of integers is $\mathbb{Z}[\zeta_{\mathfrak{f}}]$, also called the *cyclotomic ring*. In this case, there are $2s_2 = n = \varphi(\mathfrak{f})$ complex canonical embeddings (no real ones), defined by $\sigma_i(\zeta_{\mathfrak{f}}) = \zeta_{\mathfrak{f}}^i$ for $i \in \mathbb{Z}_{\mathfrak{f}}^*$. For an element $\mathbf{x} = \zeta^j \in K$ in the power basis of K, all the embeddings of $\mathbf{x}$ have magnitude 1, and hence $\|\mathbf{x}\|_2^{can} = \sqrt{n}$ and $\|\mathbf{x}\|_\infty^{can} = 1$ as well as the coefficient embedding. The discriminant of the $\mathfrak{f}$th cyclotomic field of degree $n = \varphi(\mathfrak{f})$ is $\Delta_K \leq n^n$.

In the cyclotomic case, we can define the maximal real subfield $K^+ = \mathbb{Q}[\zeta_{\mathfrak{f}} + \zeta_{\mathfrak{f}}^{-1}]$, which only contains real numbers. It has index 2 in K and its degree is $n/2$. The rings of integers $\mathcal{O}_{K^+}$ of K^+ is simply $\mathbb{Z}[\zeta_{\mathfrak{f}} + \zeta_{\mathfrak{f}}^{-1}]$. The embeddings σ_1, σ_{-1} both fix every elements in $\mathbb{K}^+$ and the relative norm $\mathsf{N}_{K/K^+}(\mathbf{a}) = \sigma_1(\mathbf{a}) \cdot$

$\sigma_{-1}(\mathbf{a}) = \mathbf{a}\cdot\bar{\mathbf{a}}$. If we represent $\mathbf{a}$ as a polynomial $\mathbf{a}(\mathbf{x}) = \sum_{i=0}^{n-1} \mathbf{a}_i\mathbf{x}^i \in \mathbb{Q}[\mathbf{x}]/\Phi_\mathfrak{f}(\mathbf{x})$, then $\bar{\mathbf{a}}(\mathbf{x}) = \mathbf{a}(1/\mathbf{x}) = a_0 - \sum_{i=1}^{n-1} a_i\mathbf{x}^i$.

Ideals in the Ring of Integers. The ring of integers $\mathcal{O}_K$ of a number field K of degree n is a free $\mathbb{Z}$-module of rank n, i.e. the set of all $\mathbb{Z}$-linear combinations of some *integral basis* $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\} \subset \mathcal{O}_K$. It is also a $\mathbb{Q}$-basis for K. In the case of cyclotomic field, the power basis $\{1, \zeta_\mathfrak{f}, \ldots, \zeta_\mathfrak{f}^{n-1}\}$ is an integral basis of the cyclotomic ring $\mathbb{Z}[\zeta_\mathfrak{f}]$ which is isomorphic to $\mathbb{Z}^n$ with $n = \varphi(\mathfrak{f})$.

It is well known that

$$\mathrm{Vol}(\mathbb{Z}[\zeta_\mathfrak{f}])^2 = \frac{\mathfrak{f}^{\phi(\mathfrak{f})}}{\prod_{p|\mathfrak{f}} p^{\phi(\mathfrak{f})/(p-1)}}.$$

In particular, if $\mathfrak{f}$ is a power of two, $\mathrm{Vol}(\mathbb{Z}[\zeta_\mathfrak{f}]) = (\mathfrak{f}/2)^{\mathfrak{f}/4}$. In this case, we also have that $(\zeta_\mathfrak{f}^i)_{i=0}^{\mathfrak{f}/2-1}$ is an orthogonal basis for the norm $\|\cdot\|$.

Lattices. Let $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis of a lattice $\mathcal{L}$. Given $\mathbf{B}$, the LLL algorithm outputs a vector $\mathbf{v} \in \mathcal{L}$ satisfying $\|\mathbf{v}\|_2 \leq 2^{n/2}\cdot\det(\mathcal{L})^{1/n}$ in polynomial time in the size of its input.

Theorem 1. ***(Minkowski)*** *For any lattice $\mathcal{L}$ of dimension n, there exists $\mathbf{x} \in \mathcal{L} \setminus \{\mathbf{0}\}$ with $\|\mathbf{x}\| \leq \sqrt{n}\mathrm{Vol}(\mathcal{L})^{1/n}$.*

We give a theorem for estimating the running time of lattice based algorithms:

Theorem 2. *Given a lattice $\mathcal{L}$ of dimension n, we can find a non-zero vector in $\mathcal{L}$ of norm less than $\beta^{n/\beta}\mathrm{Vol}(\mathcal{L})^{1/n}$ in deterministic time smaller than $2^{O(\beta)}$ times the size of the description of $\mathcal{L}$, for any $\beta < n/2$. With b_i^* the Gram-Schmidt norms of the output basis, we have $b_i^*/b_j^* \leq \beta^{O((j-i)/\beta+\log\beta)}$. Furthermore, the maximum of the Gram-Schmidt norms of the output basis is at most the maximum of the Gram-Schmidt norms of the input basis.*

Proof. Combine the semi-block Korkin-Zolotarev reduction [41] and the efficient deterministic shortest vector algorithm [36] with block size $\Theta(\beta)$ for the first point. Schnorr's algorithm combines the use of LLL reduction on a (possibly) linearly dependent basis, which is known to not increase the maximum of the Gram-Schmidt norms, and the insertion of a vector in position i whose projected norm is less than b_i^*. Also, the b_i^* decrease by a factor of at most $\beta^{O(\log\beta)}$ in a block, and the first Gram-Schmidt norms of blocks decrease by a factor of at most $\beta^{O(\beta)}$. □

Lattice Analysis. We also use the GSA assumption [42], which states that the Gram-Schmidt norms output by a lattice reduction follow a geometric sequence. If we draw the the curve with the log of the Gram-Schmidt norms, we see a line with slope $\log\beta/\beta$ is the case of BKW (it is not accurate for the last ones than follows a parabola instead). Usually, we use the fact that the minimum of the Gram-Schmidt norms has to be smaller than the norm of the smallest vector in

order to find it and so, the slope has to be close to horizontal, which implies that β is large.

In our analysis, we will a result of Pataki and Tural [39] in order to take into account that in NTRU lattice, all the shifts form a sublattice with small volume. They proved that the volume of the sublattice generated by r vectors is larger than the product of the r smallest Gram-Schmidt norms.

Lemma 1 ([39]). *Let $\mathcal{L} \subseteq \mathbb{R}^n$ be a full-rank lattice and $r \geq 1$. Then for any basis $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of $\mathcal{L}$, and any r-dimensional sublattice $\mathcal{L}'$ of $\mathcal{L}$, we have*

$$\det(\mathcal{L}') \geq \min_{1 \leq t_1 < \cdots < t_r \leq n} \prod_{1 \leq i \leq r} b^*_{t_i}.$$

Distribution on Ideal Lattices. The discrete Gaussian distribution over a lattice $\mathcal{L}$ is noted $D_{\mathcal{L},s}$, where the probability of sampling $\mathbf{x} \in \mathcal{L}$ is proportional to $\exp(-\pi\|\mathbf{x}\|^2/s^2)$. The continuous Gaussian distribution over K is noted D_s, and its density in $\mathbf{x}$ is proportional to $\exp(-\pi\|\mathbf{x}\|^2/s^2)$. We define

$$\rho_s(E) = \sum_{\mathbf{x} \in E} \exp(-\pi\|\mathbf{x}\|^2/s^2).$$

We will denote by $\mathbb{E}[X]$ the expectation of a random variable X.

We now recall two results from [35] and Banaszczyk's lemma [4] about discrete gaussian sampling over a lattice.

Lemma 2. *Given a lattice $\mathcal{L} \subset \mathbb{R}^n$, for any s and $\mathbf{c} \in \mathbb{R}^n$, we have*

$$\rho_s(\mathcal{L} + \mathbf{c}) \leq \rho_s(\mathcal{L}).$$

Lemma 3. *For a lattice $\mathcal{L}$, any $t \geq 1$, the probability that $\mathbf{x}$ sampled according to $D_{\Lambda,s}$ verifies $\|\mathbf{x}\| > st\sqrt{\frac{n}{2\pi}}$ is at most*

$$\exp\big(-n(t-1)^2/2\big).$$

We now prove a standard bound on ideal lattices, which indicates that they do not have very short vectors :

Lemma 4. *Let $M \subset (K \otimes \mathbb{R})^d$ be an $\mathcal{O}$ module of rank 1. Then, for any $0 \neq \mathbf{v} \in M$, we have $\mathrm{Vol}(M) \leq \sqrt{\Delta}\|\mathbf{v}/\sqrt{n}\|^n$.*

Proof. Since we can build a K-linear isometry from $\mathbb{R} \otimes M$ to $K \otimes \mathbb{R}$, we can assume $d = 1$. Then,

$$\mathrm{Vol}(M) \leq \mathrm{Vol}(\mathbf{v}\mathcal{O}) = \mathsf{N}_{K/\mathbb{Q}}(\mathbf{v})\sqrt{\Delta} \leq \|\mathbf{v}/\sqrt{n}\|^n\sqrt{\Delta}.$$

□

3 Projection of a Subring Attack

In this section, we propose a new subfield attack, that we call subring, since we use the multiplication by the original public key $\mathbf{h}$, which is an element of the n-dimensional ring R_q, in a subring for instance the maximal real ring of integers $\mathbb{Z}[X + 1/X]$ of dimension $n/2$, or in a smaller subring. First, we first show that small vectors in this lattice are linked to the norms and in the case of the maximal real ring, the short vector is $(\mathbf{f}\bar{\mathbf{g}}, \mathbf{g}\bar{\mathbf{g}})$. For some parameters, we also show that the norm is not the smallest element: this explains some experiments in [1]. Then, we show that in the case of power of two cyclotomic fields, if we project the matrix represented the subring lattice on the last d rows and columns, we can precisely analyze the running time of the algorithm. Moreover, removing some rows allows to reach optimal parameters for our subring attack, which is not possible in other subfield attacks.

3.1 Description of the basic subring attack

We show that in our subring attack, the lattice vector we are looking for is short. We first make sure that $\mathcal{O}$ is stable by all elements of H. This can be done by computing the Hermite normal form of the concatenation of the basis of $\sigma(\mathcal{O})$ for all $\sigma \in H$. We may then call $\mathcal{O}$ the order generated by this matrix.

The attack consists in finding short vectors of the lattice generated by

$$\mathbf{A} = \begin{pmatrix} q\mathbf{I}_n & \mathbf{M}_{\mathbf{h}}^{\mathcal{O}_{\mathbb{L}}} \\ \mathbf{0} & \mathbf{I}_m \end{pmatrix}$$

by using lattice reduction. We recall that $\mathbf{h}$ is the public key, so that a basis of this lattice can be built. We want to show that $\begin{pmatrix} \mathbf{f}\mathsf{N}_{K/L}(\mathbf{g})/\mathbf{g} \\ \mathsf{N}_{K/L}(\mathbf{g}) \end{pmatrix}$ is a short vector of this lattice.

The quadratic form we reduce is actually the one induced by $\|\cdot\|$, i.e. $\|(\mathbf{x}, \mathbf{y})\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$, on this lattice.

Lemma 5. *For any $\mathbf{g} \in \mathcal{O}$, we have*

$$\mathsf{N}_{K/L}(\mathbf{g}) \in \mathbf{g}\mathcal{O} \cap \mathcal{O}_{\mathbb{L}}.$$

Proof. We have

$$\mathsf{N}_{K/L}(\mathbf{g}) = \mathbf{g} \prod_{\sigma \in H - \{1\}} \sigma(\mathbf{g})$$

so that $\mathsf{N}_{K/L}(\mathbf{g}) \in \mathbf{g}\mathcal{O}$. By definition of $\mathsf{N}_{K/L}$, we have $\mathsf{N}_{K/L}(\mathbf{g}) \in \mathbb{L}$. Therefore, $\mathsf{N}_{K/L}(\mathbf{g}) \in \mathbf{g}\mathcal{O} \cap \mathcal{O}_{\mathbb{L}}$. □

Using Banaszczyk's lemma, we will now show that integers sampled from a discrete Gaussian distribution behaves in a way similar to a continuous Gaussian distribution.

Lemma 6. *Let* $\mathbf{x}$ *be sampled according to* $D_{\mathcal{O},s}$*. Then, the probability that*

$$\|\mathbf{x}\|_{op} \geq s\sqrt{2\ln(2n/\epsilon)/\pi}$$

is at most ϵ*.*

Proof. Let $\mathbf{u}$ be a unit vector, i.e. $\|\mathbf{u}\| = 1$. Then,

$$\begin{aligned}\rho_s(\mathcal{O})\mathbb{E}[\exp(2\pi t\langle \mathbf{x}, \mathbf{u}\rangle/s^2)] &= \sum_{\mathbf{x}\in\mathcal{O}} \exp(-\pi(\langle \mathbf{x}, \mathbf{x}\rangle - 2\langle \mathbf{x}, t\mathbf{u}\rangle)/s^2) \\ &= \exp(\pi t^2/s^2)\sum_{\mathbf{x}\in\mathcal{O}} \exp(-\pi\|\mathbf{x} - t\mathbf{u}\|^2/s^2) \\ &= \exp(\pi t^2/s^2)\rho_s(\mathcal{O} - t\mathbf{u}).\end{aligned}$$

We deduce with the previous lemma

$$\mathbb{E}[\exp(2\pi t\langle \mathbf{x}, \mathbf{u}\rangle/s^2)] \leq \exp(\pi t^2/s^2).$$

Using Markov's inequality and the union bound with $-\mathbf{u}$, we have that the probability of $|\langle \mathbf{x}, \mathbf{u}\rangle| \geq t$ is at most $2\exp(-\pi t^2/s^2)$.

We now use $t = s\sqrt{\ln(2n/\epsilon)/\pi}$, so that the probability of any real or imaginary part of a coordinate of $\mathbf{x}$ in $\mathbb{R}^r\mathbb{C}^s$ is larger than

$$s\sqrt{\ln(2n/\epsilon)/\pi}$$

is at most ϵ. □

Theorem 3. *Let* $\mathbf{f}$ *be sampled according to* $D_{\mathcal{O},\sigma}$*,* $\mathbf{g}$ *according to* $D_{\mathcal{O},s}$ *and set* $\mathbf{h} = \mathbf{f}/\mathbf{g}$*. Assume* $\mathbf{h}$ *is well defined, except with probability at most* $\epsilon/3$*. Then, there exists* $\mathbf{x} \neq \mathbf{0}$ *where* $\mathbf{x}$ *is an integer vector, such that*

$$\|\mathbf{A}\mathbf{x}\| \leq \sqrt{n(1 + \sigma^2/s^2)}\left(s\sqrt{2\ln(6n/\epsilon)/\pi}\right)^{n/m}$$

except with probability at most ϵ*.*

Proof. With probability at least $1 - \epsilon$, we have

$$\|\mathbf{f}\|_{op} \leq \sigma\sqrt{2\ln(6n/\epsilon)/\pi}$$

and

$$\|\mathbf{g}\|_{op} \leq s\sqrt{2\ln(6n/\epsilon)/\pi}.$$

In this case, we consider $\mathbf{y}$ such that $\mathbf{h}\mathsf{N}_{K/L}(\mathbf{g}) + q\mathbf{y} = \mathbf{f}\mathsf{N}_{K/L}(\mathbf{g})/\mathbf{g}$ and consider

$$\mathbf{x} = \begin{pmatrix} \mathbf{y} \\ \mathsf{N}_{K/L}(\mathbf{g}) \end{pmatrix}.$$

Using the multiplicativity of operator norms, we have

$$\|\mathsf{N}_{K/L}(\mathbf{g})\|_{op} \leq \left(s\sqrt{2\ln(6n/\epsilon)/\pi}\right)^{|H|}$$

and

$$\|\mathbf{f}\mathsf{N}_{K/L}(\mathbf{g})/\mathbf{g}\|_{op} \leq \sigma/s\left(s\sqrt{2\ln(6n/\epsilon)/\pi}\right)^{|H|}.$$

Finally,

$$\|\mathbf{Ax}\|^2 = \|\mathbf{f}\mathsf{N}_{K/L}(\mathbf{g})/\mathbf{g}\|^2 + \|\mathsf{N}_{K/L}(\mathbf{g})\|^2 \leq n\big(\|\mathbf{f}\mathsf{N}_{K/L}(\mathbf{g})/\mathbf{g}\|_{op}^2 + \|\mathsf{N}_{K/L}(\mathbf{g})\|_{op}^2\big).$$

□

We now try to get rid of the factor $\Theta(\ln(6n/\epsilon))^{n/2m}$ which is significant when s is small and n/m is large. To do so, we *h*euristically assume that $D_{\mathcal{O},\sigma}$ has properties similar to a *continuous* Gaussian here.

Theorem 4. *Let* $\mathbf{f}$ *be sampled according to* D_s *and* $E \subset G$*. Then, except with probability at most* ϵ *and under heuristics, we have :*

$$\left\| \prod_{\sigma\in E} \sigma(\mathbf{f}) \right\|_{op} \leq \Theta(s)^{|E|} \exp\left(\Theta(\sqrt{|E|\log(n/\epsilon)}\right)$$

under the condition $|E| = \Omega\big(\log(n/\epsilon)\log^2(\log(n/\epsilon))\big)$.

Proof. Let X be a random variable over $\mathbb{R}^+$, with a probability density function proportional to $\exp(-\pi x^2/s^2)$; and $Y = \sqrt{X_0^2 + X_1^2}$ where X_0 and X_1 are independent copies of X.

We have $\mathbb{E}[\log(X)] = \log(s) + \Theta(1)$ and $\mathrm{Var}[\log(X)] = \Theta(1)$ and $\log(X) < \log(s)+\Theta(\log(\log(n/\epsilon)))$ except with probability $\epsilon/(2n^2)$, due to standard bounds on Gaussian tails. Also, the same is true for Y.

We can now use the one-sided version of Bernstein's inequality [8, Theorem 3] : for Z the average of $|E|$ independent copies of $\log(X)$ or $\log(Y)$, we have :

$$\Pr[Z > t + \log(s)] \leq \epsilon/(2n) + \exp\left(-\frac{|E|t^2}{2(\Theta(1) + \Theta(\log(\log(n/\epsilon)))t/3)}\right).$$

We then choose some $t = \Theta\big(\sqrt{\log(n/\epsilon)/|E|}\big)$, so that with our lower bound on $|E|$, this probability is at most ϵ/n.

The result follows from the union bound over the coordinates in the canonical embedding of $\prod_{\sigma\in E}\sigma(\mathbf{f})$. □

For some parameters, the norm may not be the shortest element, as demonstrated by the following theorem.

Theorem 5. *There exists an element* $\mathbf{v} \in \mathbf{g}\mathcal{O} \cap \mathcal{O}_{\mathbb{L}}$ *with*

$$0 < \|\mathbf{v}\| \leq \sqrt{m}\Delta^{1/(2n)}\sigma^{n/m}$$

with probability $1 - 2^{-\Omega(n)}$.

Proof. We use Banaszczyk's lemma with $t = 2$, so that $\|\mathbf{g}\| \leq \sigma\sqrt{2n/\pi}$ except with probability $\exp(-n/2)$. Then, the determinant of $\mathbf{v} \in \mathbf{g}\mathcal{O} \cap \mathcal{O}_{\mathbb{L}}$ is smaller than the determinant of $\mathsf{N}_{K/L}(\mathbf{g})\mathcal{O}_{\mathbb{L}}$, which is $\mathsf{N}_{K/\mathbb{Q}}(\mathbf{g})\sqrt{\Delta_{\mathbb{L}}}$. But we have $\mathsf{N}_{K/\mathbb{Q}}(\mathbf{g}) \leq \left(\frac{\|\mathbf{g}\|}{\sqrt{n}}\right)^n$ and $\Delta_{\mathbb{L}} \leq \Delta^{m/n}$ so we conclude with Minkowski's theorem. □

This implies that for most parameters, the norm of the shortest non-zero vector is around $O(\sigma)^{n/m}$. Since this is smaller than the previous value as soon as n/m is a bit large, it explains why [1] found vectors shorter than the solution.

3.2 Asymptotic analysis for power of two cyclotomic fields

We set here $\mathbb{K} = \mathbb{Q}[X]/(X^n + 1) \simeq \mathbb{Q}[\zeta_{2n}]$ for n a power of two, and $\mathcal{O} = \mathbb{Z}[X]/(X^n + 1) \simeq \mathbb{Z}[\zeta_{2n}]$ which is popular in cryptosystems. For some $r \mid n$ (any such r works), we select $\mathbb{L} = \mathbb{Q}[X^r]$ so that $\mathcal{O}_{\mathbb{L}} = \mathbb{Z}[X^r]$ and $|H| = r$, so that m, the dimension of $\mathbb{L}$ is n/r. Since the $\mathbf{X}^i$ forms an orthogonal basis, we have that the coordinates of $\mathbf{f}$ and $\mathbf{g}$ are independent discrete Gaussians of parameter $s/\sqrt{n}$. Also, we can directly reduce the lattice generated by $\mathbf{A}$ with the canonical quadratic form.

We restrict our study to power of two cyclotomic fields because $\mathcal{O}$ has a known orthogonal basis, so that we can derive a closed-form expression of the results. In more complicated cases, it is clear that we can deduce the result using a polynomial time algorithm.

For the rest of this section, we assume that when the previous algorithm is used on our orthogonal projection of $\mathbf{A}\mathbb{Z}^{n+m}$, and finds a vector shorter than $\sqrt{k}\mathrm{Vol}(\mathcal{L})^{1/k}$ (which is about the size of the shortest vector of a random lattice if the lattice dimension is k), then it must be a short multiple of the key. This assumption is backed by all experiments in the literature, including ours, and can be justified by the fact that decisional problems over lattices are usually as hard as their search counterpart (see [34] for example).

We also assume the size of the input is in $n^{O(1)}$, which is the usual case.

Theorem 6. *Let* $nB^2 = \|\mathsf{f}\mathsf{N}_{K/L}(\mathbf{g})/\mathbf{g}\|^2 + \|\mathsf{N}_{K/L}(\mathbf{g})\|^2$. *Assume* $\frac{\log(qB)}{\log(q/B)} \leq r$. *Then, for*

$$\frac{\beta}{\log \beta} = \frac{2m \log q}{\log(q/B)^2}$$

we can find a non-zero element $\mathbf{Ax}$ *such that* $\|\mathbf{Ax}\|^2 = O(nB^2)$ *in time* $2^{O(\beta + \log n)}$.

Proof. We extract the last $d \approx m\frac{\log(q^2)}{\log(q/B)} \leq n+m$ rows and columns of

$$\mathbf{A} = \begin{pmatrix} q\mathbf{I} & \mathbf{M}_{\mathbf{h}}^{\mathcal{O}_{\mathbb{L}}} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

and call the generated lattice $\mathcal{L}$. Note that it is the lattice generated by $\mathbf{A}$ projected orthogonally to the first columns, so that it contains a non-zero vector $\mathbf{y}$ such that $\|\mathbf{y}\|^2 \leq nB^2$. Then, we can compute the needed β by

$$\begin{aligned}
\frac{1}{d}\log\left(\frac{\sqrt{n}\mathrm{Vol}(\mathcal{L})^{1/d}}{\sqrt{n}B}\right) =& \frac{d-m}{d^2}\log(q) - \frac{1}{d}\log(B) \\
\approx& \frac{\log(q/B)}{m\log(q^2)}\left(\frac{\log(qB)\log(q/B)\log(q)}{\log(q/B)\log(q^2)} - \log(B)\right) \\
=& \frac{\log(q/B)}{m\log(q^2)}\left(\frac{\log(qB)}{2} - \log(B)\right) \\
=& \frac{\log^2(q/B)}{2m\log(q)}.
\end{aligned}$$

The previous theorem indicates we can recover a short vector $\mathbf{z} \neq \mathbf{0}$ in $\mathcal{L}$ with $\|\mathbf{z}\| \leq nB^2$ in time $2^{\Theta(\beta+\log n)}$, and our assumption implies it is in fact a short vector in $\mathbf{A}\mathbb{Z}^{n+m}$. □

Notice that for $B \leq q$, a necessary condition for the problem to be solvable, we have $d \geq 2m$. It implies that the optimal dimension d cannot be reached by previous algorithms.

Theorem 7. *Let* $\mathbf{f}$ *and* $\mathbf{g}$ *be sampled according to* $D_{\mathcal{O},\sigma}$*, and* $\mathbf{h} = \mathbf{f}/\mathbf{g}$ *mod* q *which is well defined with probability at least* $1-\epsilon$*. Assume* $\sigma = n^{\Omega(1)}$ *and* $\sigma < q^{1/4}$*. Then, we can recover a non-zero multiple of* $(\mathbf{f}, \mathbf{g})$ *of norm at most* $\sqrt{q}$ *in time*

$$\exp\left(O\left(\max\left(\log n, \frac{n\log\sigma}{\log^2 q}\log\left(\frac{n\log\sigma}{\log^2 q}\right)\right)\right)\right)$$

with a probability of failure of at most $\epsilon + 2^{-n}$*.*

This is polynomial time for

$$\log\sigma = O\left(\frac{\log^2 q \log n}{n\log\log n}\right).$$

Proof. We choose $m = \Theta(\max(1, \frac{n\log\sigma}{\log q})) \leq n$ so that we can set $B = \sqrt{q}$, except with probability ϵ. The corresponding β is given by

$$\frac{\beta}{\log\beta} = \frac{2m\log q}{\log(q/B)^2} = \Theta(m/\log(q)) = \Theta\left(\frac{n\log\sigma}{\log^2 q}\right).$$

□

If we use $\log \sigma = \Theta(\log n)$ as in many applications, we are in polynomial time when

$$q = 2^{\Omega(\sqrt{n \log \log n})}.$$

If $\sigma = \Theta(\sqrt{n})$, the best generic algorithm runs in time $2^{\Theta(n/\log \log q)}$, which is slower for any $q \geq n^{\Theta(\sqrt{\log \log n})}$.

3.3 Comparison with other subfield attacks

We consider the lattice generated by $\begin{pmatrix} q\mathbf{I}_n & \mathbf{M}_{\mathbf{h}}^{\mathcal{O}_{\mathbb{L}}} \\ \mathbf{0} & \mathbf{I}_{n/r} \end{pmatrix}$ while Albrecht *et al.* for instance consider $\begin{pmatrix} q\mathbf{I}_{n/r} & \mathbf{M}_{\mathsf{N}_{K/L}(\mathbf{h})}^{\mathcal{O}_{\mathbb{L}}} \\ \mathbf{0} & \mathbf{I}_{n/r} \end{pmatrix}$, where $\mathbf{M}_{\mathbf{h}}^{\mathcal{O}_{\mathbb{L}}}$ represents the multiplication by the element $\mathbf{h}$ in the subring $\mathcal{O}_{\mathbb{L}}$ of $\mathbb{K}$. Our lattice is of dimension $n+n/r$, which is *larger* than Albrecht *et al.* attack, but smaller than the $2n$ original lattice. Since the running time of lattice reduction algorithms depends on the dimension of the matrix, we may think that our variant is less efficient than the subfield attack. First of all, in order to improve the running time, we will show that we can work in a projected lattice and not on the full $(n + n/r, n + n/r)$-matrix by considering the matrix forms using the last d rows and columns. The idea of working in this lattice is that the second important parameter is the *approximation factor*. This parameter depends on the size of the Gram-Schmidt coefficients. If we use the logarithm of their size, these coefficients draw a decreasing line of slope correlated with the approximation factor, so that the smaller the approximation factor be, the more horizontal the line will be. However, if we have only a $(2n/r)$-dimensional matrix, as in the subfield attack, *the determinant is too small to produce large Gram-Schmidt norms*. This problem is bypassed with our approach since we can choose the number of coordinates and the size of the subfield. Using this attack, we were able to break in practice proposed parameters by YASHE and in other papers, which were not the case in Albrecht *et al.*. We also show a tight estimation of the parameters broken by lattice reduction, and in particular that working in the original field works well. Experiments were conducted in an extensive way, and over much larger parameters.

4 Analysis of lattice reduction on NTRU lattices

We now show how to predict when this attack will work, and compare our theoretical analysis with experiments. While Albrecht *et al.* compare the subfield attack to the attack on the full dimension lattice, we will show that, the classical attack, used in Howgrave-Graham work on the hybrid attack, performs a lattice reduction on the matrix centered in the original Coppersmith-Shamir lattice. This gives a better result and we show that considering subfield is not helpful. Consequently, this attack can also be mounted on NTRU prime with overstretched parameters and works well.

4.1 Analysis of the simple method

Here, we consider the lattice reduction algorithm described in Theorem 2 applied to the full Coppersmith-Shamir matrix. We show that using Pataki-Tural lemma and the above heuristics, we can actually achieve the same efficiency regardless of the presence of a subfield, as long as we know an orthogonal basis of $\mathcal{O}$.

The analysis hinges on the fact that the difficulty for lattice reduction to find a vector in a sublattice of low volume depends on the rank of the sublattice. Previous analysis relied on its special case where the rank is one, so that the volume is the length of the generator. In this case, we can prove using the GSA and the determinant of the lattice, that $\beta/\log\beta = O\big(n/\log(q/\sigma^2)\big)$. In the following, using the Pataki-Tural lemma, we show that we can achieve the same efficiency as in the case of subfield, directly on the Coppersmith-Shamir lattice, *i.e.* $\beta/\log\beta = O\big(n\log(\sigma)/\log^2(q)\big)$.

The following theorem, identical to [1, Theorem 2], indicates that short vectors are multiples of the secret key.

Theorem 8. *Let* $\mathbf{f}, \mathbf{g} \in \mathcal{O}$ *with* $\mathbf{g}$ *invertible modulo* q *and* $\mathbf{f}$ *coprime to* $\mathbf{g}$*. Then, any vector shorter than* $\frac{nq}{\|(\mathbf{f},\mathbf{g})\|}$ *in*

$$\begin{pmatrix} q\mathbf{I}_n & \mathbf{M}^{\mathcal{O}}_{\mathbf{f}/\mathbf{g}} \\ 0 & \mathbf{I}_n \end{pmatrix} \mathcal{O}^2 \text{ is in } \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix} \mathcal{O}.$$

Proof. By coprimality, there exists $\mathbf{F}, \mathbf{G}$ such that $\mathbf{fG} - \mathbf{gF} = q$. Then,

$$\begin{pmatrix} \mathbf{f} & \mathbf{F} \\ \mathbf{g} & \mathbf{G} \end{pmatrix}$$

generates the same lattice. We let $\Lambda = \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix} \mathcal{O} \subset (\mathbb{R} \otimes \mathbb{K})^2$ and Λ^* the projection of $\begin{pmatrix} \mathbf{F} \\ \mathbf{G} \end{pmatrix} \mathcal{O}$ orthogonally to Λ. We have $\mathrm{Vol}(\Lambda)\mathrm{Vol}(\Lambda^*) = q^n\Delta$. Finally, let $0 \neq \mathbf{x} \in \Lambda^*$. Using twice section 4, we have

$$\|\mathbf{x}/\sqrt{n}\|^n \geq \frac{q^n\Delta}{\sqrt{\Delta}\mathrm{Vol}(\Lambda^*)} = \big(\frac{q\sqrt{n}}{\|(\mathbf{f},\mathbf{g})\|}\big)^n.$$

□

In the following, we show that the Pataki-Tural lemma allows us to have a lattice reduction algorithm with β around $\tilde{\Theta}(n\log\sigma/\log^2 q)$, which is close to theorem 7 in the case of subfield.

Theorem 9. *Let* $\mathbf{f}, \mathbf{g}$ *sampled according* $D_{\mathcal{O},s}$ *such that* $\mathbf{g}$ *is invertible with probability* $1-\epsilon$*, and an orthogonal basis of* $\mathcal{O}$ *is known. Reducing the lattice generated*

by $\begin{pmatrix} q\mathbf{I}_n & \mathbf{M}_{\mathbf{h}}^{\mathcal{O}} \\ \mathbf{0} & \mathbf{I}_n \end{pmatrix}$ *using the algorithm of section 2, assuming the minimum of the Gram-Schmidt norms does not decrease, with*

$$\beta = \Theta\left(\frac{n \log \sigma}{\log^2 q} \log\left(\frac{n \log \sigma}{\log^2 q}\right)\right),$$

we recover at least $n/2$ *vectors of a basis of* $\mathbf{f}\mathcal{O}$ *and* $\mathbf{g}\mathcal{O}$*, if* $\Delta^{1/2n}\sigma = q^{O(1)}$ *and*

$$\log q = \Omega\left(\log^2\left(\frac{n \log \sigma}{\log q}\right)\right),$$

with probability $1 - \epsilon - 2^{-\Omega(n)}$.

Proof. Before calling the lattice reduction algorithm, the Gram-Schmidt norms are $q\Delta^{1/2n}$ for the first n vectors, and $\Delta^{1/2n}$ for the next n vectors. The lattice contains $\begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix} \mathcal{O}$ so that the lattice spanned has a volume of $\sigma^n\sqrt{\Delta}$ except with probability $2^{-\Omega(n)}$, thanks to section 3.

We consider now the $2n$-dimensional basis outputted by the reduction algorithm (section 2), and call b_i^* 'small' when it is amongst the n smallest Gram-Schmidt norms, and 'large' otherwise. Let $\ell = O\left(\frac{n \log \sigma}{\log q}\right) \leq n$.

We consider two cases, depending whether there is a small b_i^* that has a large value or not. We show that either case is is impossible, which will complete the argument by contradiction. Assume first that there is an $i \leq n$, such that

$$b_i^* \geq \frac{\sqrt{n}q}{\sigma} \geq q^{1/4}\Delta^{1/2n}.$$

Suppose then again by contradiction, that there is a $b_j^* \geq q^{1/4}\Delta^{1/2n}$ which is small (Case 1). Consequently by Theorem 2,

$$b_k^* \geq q^{1/4}\Delta^{1/2n}\beta^{-O(\ell/\beta + \log \beta)} \geq q^{1/4}\Delta^{1/2n}\beta^{-O(\ell/\beta)}$$

for all the ℓ first $k \geq i$ such that b_k^* is small (we use here the assumption that the minimum of the Gram-Schmidt norms does not decrease). Hence, the product of the n smallest b_i^* is at least $\Delta^{1/2}q^{\ell/4}\beta^{-O(\ell^2/\beta)}$ by lower bounding the last $(n-\ell)$ ones by $\Delta^{1/2n}$. We deduce that for small enough constants, this is impossible using the Pataki-Tural lemma: otherwise we get a contradiction with the fact that this product should be smaller than the smallest volume of a sublattice of dimension n, $\sqrt{\Delta}\sigma^n$.

Suppose now every small b_j^* satisfies $b_j^* < q(1/4)\Delta^{1/(2n)}$ (Case 2). Let $j \geq i$ be the smallest such that b_j^* is small. Then, we have

$$b_k^* \leq q^{1/4}\Delta^{1/2n}\beta^{O(\ell/\beta + \log \beta)}$$

for all the last ℓ indices $k \leq j$ such that b_k^* is large. Thus, the product of the large Gram-Schmidt norms is at most $\Delta^{1/2}q^{n-\ell/4}\beta^{O(\ell^2/\beta)}$, as all b_k^*'s remain

$\leq q\Delta^{1/(2n)}$, by Theorem 2. Since the determinant is preserved during the running of the algorithm, the product of the small Gram-Schmidt norms is at least $\Delta^{1/2}q^{\ell/4}\beta^{-O(\ell^2/\beta)}$, which is impossible by using again the Pataki-Tural lemma.

To sum up, we have proved that for all the n first b_i^*, we have $b_i^* < \sqrt{n}q/\sigma$ and so, $b_i < nq/\sigma$ and using lemma 8, we can show that all the first $n/2$ vectors are in $\begin{pmatrix}\mathbf{f}\\ \mathbf{g}\end{pmatrix}\mathcal{O}$. □

4.2 Generalization and the middle technique

As we can see from the formula, considering a subfield is not helpful since the quantity $n\log\sigma$ is essentially constant; unless we have reasons to believe there are huge factors of $\mathbf{g}\mathcal{O}$ which are in the subfield. Even worse, it actually decreases the efficiency when $\sigma \geq \sqrt{q}$ because the value of ℓ is forced to a suboptimal. We also observe that the significant reduction in the dimension due to the use of subfields, allowing to break instances of high dimension is also present here : indeed, we can project orthogonally to the first $2n-\ell$ vectors the next ℓ vectors so that we reduce a lattice of dimension ℓ instead of $2n$.

Also, when we choose to work with $\mathcal{O} = \mathbb{Z}[X]/(X^n - X - 1)$ as in NTRU Prime [5], where we can use $(X^i)_{i=0}^{n-1}$ as an orthogonal basis due to the choice of the error distribution made by the authors (the coordinates are almost independent and uniform over $\{-1, 0, 1\}$), the same result applies.

We stress that while our theorem does not prove much - assuming the maximum of the Gram-Schmidt norms decreases is wrong, except for LLL - experiments indicate that either the middle part of the lattice behaves as a 'random' lattice as it is evaluated in [25], or the first n vectors are a basis of $\begin{pmatrix}\mathbf{f}\\ \mathbf{g}\end{pmatrix}\mathcal{O}$. Furthermore, the phase transition between the two possible outputs is almost given by the impossibility of the first case. As lattice reduction algorithms are well understood (see [24,11]), it is thus easy to compute the actual β.

5 Implementation

Heuristically, we have that for reduced random lattices, the sequence b_i^* is (mostly) decreasing and therefore the relevant quantity is $\prod_{i=n-r}^{n-1} b_i^*$. It means that when the b_i^*s decrease geometrically and $\det(\mathcal{L}')^{1/r}$ is about the length of the shortest vector, we need $b^*_{\lfloor n-r/2\rfloor}$ to be larger than the shortest vector instead of the b^*_{n-1} given by a standard analysis. We now remark that for $r = 1$, this is pessimistic. Indeed, for a "random" short vector, we expect the projection to reduce its length by a $\simeq \sqrt{n}$ factor. In our case, we can expect the projection to reduce the length by a $\simeq \sqrt{n/(n-r)}$ factor.

For our predictions, we assumed that the determinant of the quadratic form

$$\mathbf{x} \mapsto \mathbf{f}\mathsf{N}_{K/L}(\mathbf{g})/\mathbf{g}\mathbf{x}\overline{\mathbf{f}\mathsf{N}_{K/L}(\mathbf{g})/\mathbf{g}\mathbf{x}} + \mathsf{N}_{K/L}(\mathbf{g})\mathbf{x}\overline{\mathsf{N}_{K/L}(\mathbf{g})\mathbf{x}},$$

which corresponds to the $\det(\mathbf{U}^t\mathbf{G}\mathbf{U})$ above, is about the square of the norm over $\mathbb{Z}$ of $\mathbf{g}$. This quantity can be evaluated in quasi-linear time when we work within a cyclotomic field with smooth conductor by repeatedly computing the norm over a subfield, instead of the generic quadratic algorithm, or its variants such as in [2, section 5.2]. We observe a very good agreement between the experiments and the prediction, while considering only the fact that the lattice has a short vector would lead a much higher bound. Also, while $\mathsf{N}_{K/L}(\mathbf{g})$ has a predicted size of $n^{r/2}\exp(\sqrt{r\log(n/r)})$ with $\sigma = \sqrt{n}$, we expect LLL to find a multiple of size $n^{r/2}\exp(n/r)$ (possibly smaller) but none of these quantities are actually relevant for determining whether or not LLL will recover a short element.

Finally, we may have $(\mathsf{N}_{K/L}(\mathbf{g}))/((\mathbf{g})\cap\mathcal{O}_L)$ which is non-trivial. However, if it is an ideal of norm κ, we have that κ^2 divides the norm over $\mathbb{Z}$ of $\mathbf{g}$, which is exceedingly unlikely for even small values of $\kappa^{r/n}$.

Our predictions indicate all proposed parameters of [6, Table 1] are broken by LLL. We broke the first three using fplll and about three weeks of computation. The last three where broken in a few days over a 16-core processor (Intel Xeon E5-2650).

The parameters proposed for schemes using similar overstretched NTRU assumption, such as in homomorphic encryption [7,30,15,16], [14,10,31,18] or in private information retrieval [17], are also broken in practical time using LLL. For example, we recovered a decryption key of the FHE described in [15] in only 10 hours. For comparison, they evaluated AES in 29 h: that means that we can more efficiently than the FHE evalution, recover the secret, perform the AES evaluation, and then re-encrypt the result! A decryption key was recovered for [18] in 4 h. Other instanciations such as [9,28] are harder, but within range of practical cryptanalysis, using BKZ with moderate block-size [11].

6 Explicit complexity

We now turn towards the problem of deriving the first order of the asymptotical complexity of *heuristic* algorithms. Before the dual BKZ algorithm [38], simple derivations (as in [29, Appendix B]) could only be done using the Geometric Series Assumption, since the heuristic Gram-Schmidt norms outputted by the BKZ algorithm have a fairly complicated nature (see [24]), making an exact derivation quite cumbersome if not intractable. We are only interested in the part of the Gram-Schmidt norms known to be geometrically decreasing, which simplifies the computations [4].

We emphasize that we are only using standard heuristics, checked in practice, and *tight* at the first order. We compute the necessary block-size β to solve the problems and assume $\log\beta \approx \log n$. More precisely, if $\log\beta = (1+o(1))\log n$, then the exponent in the running time is within $1+o(1)$ of its actual value.

[4] We remark that the last Gram-Schmidt norms have no constraints in the original algorithm. However, we can always assume they are HKZ-reduced, so that their logarithms are a parabola.

$\log n$	$\log q$	$\log r$	Success	Method	Coordinates used	Origin
11	165	4	Yes	[1]	128	-
11	115	4	Yes	Ours	510	-
11	114	4	No	Ours	630	-
11	95	3	Yes	[1]	256	-
11	81	3	Yes	Ours	600	-
11	80	3	No	Ours	600	-
11	79	3	No	Ours	860	YASHE[6]
11	70	2	Yes	Ours	600	-
11	69	2	No	Ours	600	-
12	190	4	Yes	[1]	256	-
12	157	4	Yes	Ours	430	YASHE[6]
12	144	4	Yes	Ours	850	-
12	143	4	No	Ours	850	-
13	383	4	Yes	Ours	512	[18]
13	312	5	Yes	Ours	470	YASHE[6]
14	622	5	Yes	Ours	470	YASHE[6]
15	1271	5	Yes	Ours	512	[15]
15	1243	6	Yes	Ours	660	YASHE[6]
16	2485	7	Yes	Ours	820	YASHE[6]

$\log n$	Prediction	$\log r$
11	116	4
11	82	3
11	71	2
12	146	4
12	105	1
13	271	5
13	155	1
14	525	6
14	228	1
15	1045	7
15	335	1
16	2121	8
16	491	1

Fig. 1. Experiments with LLL for solving the NTRU problem in the ring $\mathbb{Z}[X]/(q, X^n + 1)$, where the coefficients of the polynomials are uniform in $\{-1, 0, 1\}$. The lattice dimension used is equal to the number of coordinates used added to n/r. The values of [1] are the smallest moduli for which their algorithm works, up to one, one and five bits. The prediction is the minimum $\log q$ an LLL reduction can solve assuming we use all the (necessary) coordinates.

$\log n$	$\log q$	ℓ	Success
11	72	1116	Yes
11	70	1200	Yes
11	69	1200	No
12	118	1024	Yes
12	117	1024	No
12	105	1700	Yes
12	104	1700	No
13	230	1024	Yes
14	450	1024	Yes
15	930	1024	Yes

$\log n$	ℓ	Prediction
11	1033	71
12	1472	106
13	2275	156
14	3357	230
15	5127	337
16	7124	477

Fig. 2. Experiments with LLL for solving the NTRU problem in the ring $\mathbb{Z}[X]/(q, X^p - X - 1)$, where the coefficients of the polynomials are uniform in $\{-1, 0, 1\}$ and p is the smallest prime larger than n. The lattice dimension used is ℓ. The prediction is the minimum $\log q$ an LLL reduction can solve.

For more information on the dual BKZ algorithm and dual lattices, see [38]. We denote by dual BKZ algorithm their algorithm 1 followed by a forward (i.e. primal) round, so that it attempts to *minimize* the *first* Gram-Schmidt norm (as the previous algorithms), rather than *maximizing* the *last* Gram-Schmidt norm.

We remark that all uses of NTRU for "standard" cryptography (key-exchange, signature and IBE) are instantiated with a modulus below n^2, so that the lattice reduction algorithms are *not* affected by the property.

6.1 Security of Learning With Errors

The following heuristic analysis applies for NTRU, but also for any LWE problem with dimension n and exactly $2n$ samples [5], or Ring-LWE with two samples. The primal algorithm searches for a short vector in a lattice.

As usual, we build the lattice

$$\mathbf{A} = \begin{pmatrix} q\mathbf{I}_n & \mathbf{M}_{\mathbf{h}}^{\mathcal{O}_L} \\ \mathbf{0} & \mathbf{I}_m \end{pmatrix}$$

and apply the dual BKZ algorithm on its dual. We assume it did not find the key, and suppose the projection of $(\mathbf{f}, \mathbf{g})$ orthogonally to the first $2n-1$ vector has a norm of $\sigma/\sqrt{n}$. Then, the last Gram-Schmidt norm must be smaller than $\sigma/\sqrt{n}$ and we compute the smallest block-size β such that it is not the case. Hopefully, this means that applying the algorithm with a block-size β will find the key.

Once the dual BKZ algorithm has converged, the $2n-\beta$ first Gram-Schmidt norms are decreasing with a rate of $\approx \beta^{-1/\beta}$ and the $2n-\beta$th norm is about $\sqrt{\beta}V^{1/\beta}$ where V is the product of the last β norms. We deduce that the volume of the dual lattice is

$$q^{-n} = \left(\frac{\sigma}{\sqrt{n}}\right)^{-2n} \beta^{-(2n-\beta)^2/2\beta - n} = \left(\frac{\sigma}{\sqrt{n}}\right)^{-2n} \beta^{-2n^2/\beta}$$

so with $q = n^a$, $\sigma = n^b$ and $\beta = nc$ we have

$$-a \approx 1 - 2b - 2/c$$

and we deduce $c = 2/(a+1-2b)$.

Another possibility is to apply the dual BKZ algorithm on the basis. If it reduces the last $m+n$ vectors, then the $m+n-\beta$th Gram-Schmidt norm cannot be smaller than the size of the key, σ. Now, if $m = n$ this norm is $\sqrt{q}\beta^{n/\beta-(2n-\beta)/\beta}$, and we deduce $a/2 - 1/c + 1 = b$ or $c = 2/(a+2-2b)$ which happens when $c \geq 2/a$ (iff $b \geq 1$). Else, we take m maximum so that $q^{m/(m+n)}\beta^{(m+n)/2\beta} = q$ or $m = n(\sqrt{2ca}-1)$ which gives $q\beta^{-(m+n-\beta)/\beta} = \sigma$ or $a - (\sqrt{2ca} - 1 + 1 - c)/c = b$ and hence $c = 2a/(a+1-b)^2$ when $b \leq 1$.

[5] Beware that an element sampled in the ring with standard deviation σ has coordinates of size only $\sigma/\sqrt{n}$.

The dual algorithm searches for $2^{o(n)}$ short vectors in the dual lattice, so that the inner product with a gaussian of standard deviation σ can be distinguished. Applying the dual BKZ algorithm on the dual lattice gives a vector of norm $\beta^{n/\beta} q^{-m/(n+m)} = \sigma/n$. The norm is minimized for $m = \sqrt{2ac} - 1$ or $m = n$, which gives $c = 2a/(a+1-b)^2$ when $b < 1$, and $2/(a+2-2b)$ else.

In all cases, the best complexity is given by $c = \max(2a/(a+1-b)^2, 2/(a+2-2b))$ (and when the number of samples is unlimited, this is $2a/(a+1-b)^2$).

6.2 Security of NTRU

Here, the analysis is specific to NTRU. We apply the dual BKZ algorithm to the same lattice, and compute the β such that the product of the n last Gram-Schmidt norms is equal to σ^n. Note that it is equivalent to having the product of the n first Gram-Schmidt norms equal to q/σ^n.

We first compute m such that the dual BKZ algorithm changes only the $2m$ middle norms. This is given by :

$$q = \sqrt{q}\beta^{m/\beta}$$

so that $m \approx a\beta/2$. For $a \geq 2$, we have $\beta \leq m$ so that, assuming $m \leq n$, the product of the m first norms is $q^m \beta^{-m^2/2\beta}$. Hence, we need $\beta^{m^2/2\beta} = \sigma^n$. We deduce

$$a^2c^2/8c = b$$

so that $c = 8b/a^2$.

When $m > n$, the first vector is of norm only $\sqrt{q}\beta^{n/\beta}$, so that for $c \leq 1$, we must have

$$q^{n/2}\beta^{n^2/2\beta - n^2/\beta} = \sigma^n$$

so that $a/2 - 1/2c \approx b$ and $c = 1/(a-2b)$. For this formula to be correct, we need $8b/a^2 a/2 \geq 1$, or $4b \geq a$.

We can show that this is better than the algorithms against Ring-LWE when $b = 1/2$ ($\approx$ binary errors) when $a \geq (4 + \sqrt[3]{262 - 6\sqrt{129}} + \sqrt[3]{262 + 6\sqrt{129}})/6 \approx 2.783$. When $b \geq 1$ which is the proven case, it is better for all $a > 4$ and $b < a/2 - 1$.

We again remark that going to a subfield, so that nb is constant, does not improve the complexity.

7 Conclusion

We conclude that the shortest vector problem over module lattices seems strictly easier than the bounded distance decoding. Since the practical cost of transforming a NTRU-based cryptosystem into a Ring-LWE-based cryptosystem is usually small, especially for key-exchange (e.g. [3]), we recommend to dismiss the former, in particular since it is known to be weaker (see [40, Section 4.4.4]). One important difference between NTRU and Ring-LWE instances is the fact that in

NTRU lattices, there exist many short vectors. This has been used by May and Silverman in [33] and in our case, the determinant of the sublattice generated by these short vectors is an important parameter to predict the behaviour of our algorithm.

We remark that the only proven way to use NTRU is to use $\sigma \approx \sqrt{n^3 q}$ [43]. We showed here that attacks are more efficient against NTRU than on a Ring-LWE lattice until $\sigma \approx n^{-1}\sqrt{q}$, which suggests their result is essentially optimal. Furthermore, the property we use is present until $\sigma \approx \sqrt{nq}$, i.e. until the public key $\mathbf{h}$ is (heuristically) indistinguishable from uniform.

Our results show that the root approximation factor is a poor indicator in the NTRU case : indeed, we reached 1.0059 using a mere LLL. We suggest to switch the complexity measure to the maximum dimension used in shortest vector routines (i.e. the block size of the lattice reduction algorithm) of a successful attack. While there are less problems with LWE-based cryptosystems, the root approximation factor has also several shortcomings which are corrected by this modification. Indeed, highly reduced basis do not obey to the Geometric Series Assumption, so that the root approximation factor also depends on the dimension of the lattice. Even when the dimension is much larger than the block-size, converting the factor into a block-size - which is essentially inverting the function $\beta \mapsto \left(\frac{(\beta/2)!}{\pi^{\beta/2}}\right)^{1/\beta^2}$ - is very cumbersome. Finally, the complexity of shortest vector algorithms is more naturally expressed as a function of the dimension than the asymptotical root approximation factor they can achieve.

Acknowledgments.

We also would like to thank the reviewers and particularly Damien Stehlé for his help and his advices for the final version. We would like to thank the Crypto Team at ENS for providing us computational ressources to perform our experimentations.

References

1. Martin R. Albrecht, Shi Bai, and Léo Ducas. A subfield lattice attack on overstretched NTRU assumptions - cryptanalysis of some FHE and graded encoding schemes. In Matthew Robshaw and Jonathan Katz, editors, *Advances in Cryptology – CRYPTO 2016, Part I*, volume 9814 of *Lecture Notes in Computer Science*, pages 153–178, Santa Barbara, CA, USA, August 14–18, 2016. Springer, Heidelberg, Germany.
2. Martin R. Albrecht, Catalin Cocis, Fabien Laguillaumie, and Adeline Langlois. Implementing candidate graded encoding schemes from ideal lattices. In Tetsu Iwata and Jung Hee Cheon, editors, *Advances in Cryptology – ASIACRYPT 2015, Part II*, volume 9453 of *Lecture Notes in Computer Science*, pages 752–775, Auckland, New Zealand, November 30 – December 3, 2015. Springer, Heidelberg, Germany.

3. Erdem Alkim, Léo Ducas, Thomas Pöppelmann, and Peter Schwabe. Post-quantum key exchange - a new hope. Cryptology ePrint Archive, Report 2015/1092, 2015. http://eprint.iacr.org/2015/1092.
4. Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(1):625–635, 1993.
5. Daniel J Bernstein, Chitchanok Chuengsatiansup, Tanja Lange, and Christine van Vredendaal. Ntru prime. 2016. http://eprint.iacr.org/.
6. Joppe W. Bos, Kristin Lauter, Jake Loftus, and Michael Naehrig. Improved security for a ring-based fully homomorphic encryption scheme. In Martijn Stam, editor, *14th IMA International Conference on Cryptography and Coding*, volume 8308 of *Lecture Notes in Computer Science*, pages 45–64, Oxford, UK, December 17–19, 2013. Springer, Heidelberg, Germany.
7. Joppe W Bos, Kristin Lauter, and Michael Naehrig. Private predictive analysis on encrypted medical data. *Journal of biomedical informatics*, 50:234–243, 2014.
8. Stéphane Boucheron, Gábor Lugosi, and Olivier Bousquet. Concentration inequalities. In *Advanced Lectures on Machine Learning*, pages 208–240. Springer, 2004.
9. Gizem S. Çetin, Wei Dai, Yarkın Doröz, and Berk Sunar. Homomorphic autocomplete. Cryptology ePrint Archive, Report 2015/1194, 2015. http://eprint.iacr.org/2015/1194.
10. Gizem S. Çetin, Yarkin Doröz, Berk Sunar, and Erkay Savas. Depth optimized efficient homomorphic sorting. In Kristin E. Lauter and Francisco Rodríguez-Henríquez, editors, *Progress in Cryptology - LATINCRYPT 2015: 4th International Conference on Cryptology and Information Security in Latin America*, volume 9230 of *Lecture Notes in Computer Science*, pages 61–80, Guadalajara, Mexico, August 23–26, 2015. Springer, Heidelberg, Germany.
11. Yuanmi Chen and Phong Q. Nguyen. BKZ 2.0: Better lattice security estimates. In Dong Hoon Lee and Xiaoyun Wang, editors, *Advances in Cryptology – ASIACRYPT 2011*, volume 7073 of *Lecture Notes in Computer Science*, pages 1–20, Seoul, South Korea, December 4–8, 2011. Springer, Heidelberg, Germany.
12. Jung Hee Cheon, Jinhyuck Jeong, and Changmin Lee. An Algorithm for NTRU Problems and Cryptanalysis of the GGH Multilinear Map without an encoding of zero. Cryptology ePrint Archive, Report 2016/139, 2016. http://eprint.iacr.org/.
13. Don Coppersmith and Adi Shamir. Lattice attacks on NTRU. In Walter Fumy, editor, *Advances in Cryptology – EUROCRYPT'97*, volume 1233 of *Lecture Notes in Computer Science*, pages 52–61, Konstanz, Germany, May 11–15, 1997. Springer, Heidelberg, Germany.
14. Wei Dai, Yarkin Doröz, and Berk Sunar. Accelerating SWHE based PIRs using GPUs. In Michael Brenner, Nicolas Christin, Benjamin Johnson, and Kurt Rohloff, editors, *FC 2015 Workshops*, volume 8976 of *Lecture Notes in Computer Science*, pages 160–171, San Juan, Puerto Rico, January 30, 2015. Springer, Heidelberg, Germany.
15. Yarkın Doröz, Yin Hu, and Berk Sunar. Homomorphic aes evaluation using the modified ltv scheme. *Designs, Codes and Cryptography*, pages 1–26, 2015.
16. Yarkin Doröz, Aria Shahverdi, Thomas Eisenbarth, and Berk Sunar. Toward practical homomorphic evaluation of block ciphers using prince. In Rainer Böhme, Michael Brenner, Tyler Moore, and Matthew Smith, editors, *FC 2014 Workshops*, volume 8438 of *Lecture Notes in Computer Science*, pages 208–220, Christ Church, Barbados, March 7, 2014. Springer, Heidelberg, Germany.

17. Yarkin Doröz, Berk Sunar, and Ghaith Hammouri. Bandwidth efficient PIR from NTRU. In Rainer Böhme, Michael Brenner, Tyler Moore, and Matthew Smith, editors, *FC 2014 Workshops*, volume 8438 of *Lecture Notes in Computer Science*, pages 195–207, Christ Church, Barbados, March 7, 2014. Springer, Heidelberg, Germany.
18. Nathan Dowlin, Ran Gilad-Bachrach, Kim Laine, Kristin Lauter, Michael Naehrig, and John Wernsing. Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy. 2015.
19. Léo Ducas, Alain Durmus, Tancrède Lepoint, and Vadim Lyubashevsky. Lattice signatures and bimodal Gaussians. In Ran Canetti and Juan A. Garay, editors, *Advances in Cryptology – CRYPTO 2013, Part I*, volume 8042 of *Lecture Notes in Computer Science*, pages 40–56, Santa Barbara, CA, USA, August 18–22, 2013. Springer, Heidelberg, Germany.
20. Léo Ducas, Vadim Lyubashevsky, and Thomas Prest. Efficient identity-based encryption over NTRU lattices. In Palash Sarkar and Tetsu Iwata, editors, *Advances in Cryptology – ASIACRYPT 2014, Part II*, volume 8874 of *Lecture Notes in Computer Science*, pages 22–41, Kaoshiung, Taiwan, R.O.C., December 7–11, 2014. Springer, Heidelberg, Germany.
21. Nicolas Gama, Nick Howgrave-Graham, and Phong Q. Nguyen. Symplectic lattice reduction and NTRU. In Serge Vaudenay, editor, *Advances in Cryptology – EUROCRYPT 2006*, volume 4004 of *Lecture Notes in Computer Science*, pages 233–253, St. Petersburg, Russia, May 28 – June 1, 2006. Springer, Heidelberg, Germany.
22. Craig Gentry. Key recovery and message attacks on NTRU-composite. In Birgit Pfitzmann, editor, *Advances in Cryptology – EUROCRYPT 2001*, volume 2045 of *Lecture Notes in Computer Science*, pages 182–194, Innsbruck, Austria, May 6–10, 2001. Springer, Heidelberg, Germany.
23. Craig Gentry and Michael Szydlo. Cryptanalysis of the revised NTRU signature scheme. In Lars R. Knudsen, editor, *Advances in Cryptology – EUROCRYPT 2002*, volume 2332 of *Lecture Notes in Computer Science*, pages 299–320, Amsterdam, The Netherlands, April 28 – May 2, 2002. Springer, Heidelberg, Germany.
24. Guillaume Hanrot, Xavier Pujol, and Damien Stehlé. Analyzing blockwise lattice algorithms using dynamical systems. In Phillip Rogaway, editor, *Advances in Cryptology – CRYPTO 2011*, volume 6841 of *Lecture Notes in Computer Science*, pages 447–464, Santa Barbara, CA, USA, August 14–18, 2011. Springer, Heidelberg, Germany.
25. Jeffrey Hoffstein, Jill Pipher, John M. Schanck, Joseph H. Silverman, William Whyte, and Zhenfei Zhang. Choosing parameters for ntruencrypt. In *CT-RSA*, pages 3–18, 2017.
26. Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman. NTRU: A ring-based public key cryptosystem. In *Algorithmic Number Theory, Third International Symposium, ANTS-III, Portland, Oregon, USA, June 21-25, 1998, Proceedings*, pages 267–288, 1998.
27. Nick Howgrave-Graham. A hybrid lattice-reduction and meet-in-the-middle attack against NTRU. In Alfred Menezes, editor, *Advances in Cryptology – CRYPTO 2007*, volume 4622 of *Lecture Notes in Computer Science*, pages 150–169, Santa Barbara, CA, USA, August 19–23, 2007. Springer, Heidelberg, Germany.
28. Miran Kim and Kristin Lauter. Private genome analysis through homomorphic encryption. *BMC medical informatics and decision making*, 15(Suppl 5):S3, 2015.
29. Paul Kirchner and Pierre-Alain Fouque. An improved BKW algorithm for LWE with applications to cryptography and lattices. In Rosario Gennaro and Matthew

J. B. Robshaw, editors, *Advances in Cryptology – CRYPTO 2015, Part I*, volume 9215 of *Lecture Notes in Computer Science*, pages 43–62, Santa Barbara, CA, USA, August 16–20, 2015. Springer, Heidelberg, Germany.
30. Kristin E. Lauter, Adriana López-Alt, and Michael Naehrig. Private computation on encrypted genomic data. In Diego F. Aranha and Alfred Menezes, editors, *Progress in Cryptology - LATINCRYPT 2014: 3rd International Conference on Cryptology and Information Security in Latin America*, volume 8895 of *Lecture Notes in Computer Science*, pages 3–27, Florianópolis, Brazil, September 17–19, 2015. Springer, Heidelberg, Germany.
31. Tancrède Lepoint and Michael Naehrig. A comparison of the homomorphic encryption schemes FV and YASHE. In David Pointcheval and Damien Vergnaud, editors, *AFRICACRYPT 14: 7th International Conference on Cryptology in Africa*, volume 8469 of *Lecture Notes in Computer Science*, pages 318–335, Marrakesh, Morocco, May 28–30, 2014. Springer, Heidelberg, Germany.
32. Adriana López-Alt, Eran Tromer, and Vinod Vaikuntanathan. On-the-fly multiparty computation on the cloud via multikey fully homomorphic encryption. In Howard J. Karloff and Toniann Pitassi, editors, *44th Annual ACM Symposium on Theory of Computing*, pages 1219–1234, New York, NY, USA, May 19–22, 2012. ACM Press.
33. Alexander May and Joseph H. Silverman. Dimension Reduction Methods for Convolution Modular Lattices. In *Cryptography and Lattices, International Conference, CaLC 2001, Providence, RI, USA, March 29-30, 2001, Revised Papers*, pages 110–125, 2001.
34. Daniele Micciancio and Shafi Goldwasser. *Complexity of lattice problems: a cryptographic perspective*, volume 671. Springer Science & Business Media, 2012.
35. Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on gaussian measures. *SIAM J. Comput.*, 37(1):267–302, 2007.
36. Daniele Micciancio and Panagiotis Voulgaris. Faster exponential time algorithms for the shortest vector problem. In Moses Charika, editor, *21st Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1468–1480, Austin, TX, USA, January 17–19, 2010. ACM-SIAM.
37. Daniele Micciancio and Michael Walter. Practical, predictable lattice basis reduction. In Marc Fischlin and Jean-Sébastien Coron, editors, *Advances in Cryptology – EUROCRYPT 2016, Part I*, volume 9665 of *Lecture Notes in Computer Science*, pages 820–849, Vienna, Austria, May 8–12, 2016. Springer, Heidelberg, Germany.
38. Daniele Micciancio and Michael Walter. Practical, predictable lattice basis reduction. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 820–849. Springer, 2016.
39. Gábor Pataki and Mustafa Tural. On sublattice determinants in reduced bases. *arXiv preprint arXiv:0804.4014*, 2008.
40. Chris Peikert. A decade of lattice cryptography. Cryptology ePrint Archive, Report 2015/939, 2015. `http://eprint.iacr.org/2015/939`.
41. Claus-Peter Schnorr. A hierarchy of polynomial time lattice basis reduction algorithms. *Theoretical computer science*, 53(2):201–224, 1987.
42. Claus-Peter Schnorr. Lattice reduction by random sampling and birthday methods. In *STACS*, pages 145–156, 2003.
43. Damien Stehlé and Ron Steinfeld. Making NTRU as secure as worst-case problems over ideal lattices. In Kenneth G. Paterson, editor, *Advances in Cryptology – EUROCRYPT 2011*, volume 6632 of *Lecture Notes in Computer Science*, pages 27–47, Tallinn, Estonia, May 15–19, 2011. Springer, Heidelberg, Germany.