Cryptanalysis of the Overstretched NTRU Problem for General Modulus Polynomial

Jung Hee Cheon, Minki Hhan, and Changmin Lee

Seoul National University

Abstract. The overstretched NTRU problem, which is the NTRU problem with super-polynomial size q in n, is one of the important security foundation of cryptosystems which are recently suggested. Albrecht et al. in Crypto 2016 and Cheon et al. in ANTS 2016 proposed so-called subfield attacks which demonstrate that the overstretched NTRU problems with power-of-two cyclotomic modulus are not secure enough with given parameters in GGH multilinear map and YASHE/LTV fully homomorphic encryption. Unfortunately, they heavily depend on the algebraic structure of the base ring.

On the other hand, Kirchner and Fouque presented new cryptanalysis of the overstretched NTRU problem over general modulus in Eurocrypt 2017. They achieve the similar performance compare to previous subfield attacks.

In this paper, we present a new algorithm to the overstretched NTRU problem. This algorithm has same complexity to subfield attacks, but threaten more general base ring with poly(n) expansion factor as common in suggested schemes like original GGH, YASHE scheme and NTRU prime rings. Our algorithm implies that cryptosystem related to the overstretched NTRU problem cannot be secure by changing base ring.

In addition, we present an extended (trace/norm) subfield attack for the power-of-two cyclotomic modulus. This extended subfield attack has a similar asymptotic complexity to the previous subfield attacks, but with smaller constant in the exponent term.

Keywords: NTRU, Ideal Lattice, subfield attack

1 Introduction

Given an integer polynomial $F(X)$ of degree n, a ring $R := \mathbb{Z}[X]/\langle F(X)\rangle$, a positive integer q and $\boldsymbol{h} = [\boldsymbol{f}/\boldsymbol{g}]_q \in [R]_q := \mathbb{Z}[X]/\langle F(X), q\rangle$ for $\boldsymbol{f}, \boldsymbol{g} \in R$, the NTRU problem $\mathsf{NTRU}_{R,q,M,\frac{q}{2}}$ asks to find $(\boldsymbol{d}\cdot\boldsymbol{f}, \boldsymbol{d}\cdot\boldsymbol{g})$ with bounded Euclidean norm $q/2$ for $\|\boldsymbol{f}\|, \|\boldsymbol{g}\| \leq M$.

The NTRU problem was first suggested by Hoffstein, Pipher and Silverman [HPS98] to construct a fast public key encryption scheme NTRU. More recently, this problem used as a base problem for many lattice-based cryptographic constructions including signature schemes [HHGP^{+}03,DDLL13], Fully Homomorphic Encryption (FHE) schemes [LATV12,BLLN13,DHS16], and cryptographic multilinear maps [GGH13,LSS14,ACLL14].

In the original NTRU scheme, the coefficients of $\mathbf{f}$ and $\mathbf{g}$ are restricted to $\{-1, 0, 1\}$ with small modulus q. The basic approach to solve this problem is to transform it to a short vector problem on a lattice of dimension $2n$ (called a NTRU lattice). By applying a lattice reduction algorithm (for instance, BKZ algorithm of block size β [HPS11]), it has time complexity $\mathrm{poly}(n) \cdot 2^{\Theta(\beta)}$ for $\frac{\beta}{\log \beta} = \Theta(\frac{n}{\log q})$. The best known algorithm is the combined lattice-reduction and the meet-in-the-middle attack by Howgrave-Graham [HG07] and Kirchner-Fouque algorithm [KF15]. Their time complexities are $2^{O(n)}$ and $2^{O(n/\log\log q)}$, respectively.

FHEs and multilinear maps require a larger modulus for their functionality. To distinguish it from the original NTRU, the NTRU problem with a modulus super-polynomial in n is called the *overstretched NTRU* following [ABD16]. Recently, there have been several attacks on the overstretched NTRU problem. Concurrently and independently Albrecht, Bai and Ducas [ABD16] and Cheon, Jeong and Lee [CJL16] presented two similar attacks, which are called the subfield attacks, on the overstretched NTRU problem by using the subfield structure of the base ring. They used a norm function and trace function, respectively, to reduce the dimension of the target lattice, whose basic idea can be found also in [GS02].

These algorithms have been used to cryptanalyze FHEs and multilinear maps based on NTRU-related problems: The GGH multilinear map [GGH13] over $R = \mathbb{Z}[X]/\langle X^n + 1\rangle$ with suggested parameters is broken in quasi-polynomial time of security parameter λ [ABD16,CJL16] *if n is a power of two.* The fully homomorphic encryption LTV [LATV12] and YASHE [BLLN13] are attacked in subexponential time $2^{O(\lambda/\log^{1/3}\lambda)}$ for the parameters with claimed security $2^{\Theta(\lambda)}$ [ABD16]. When the polynomial $X^n + 1$ converts to an m-th cyclotomic polynomial with a smooth integer m, the previous attacks are still applicable.

Other two approaches by Kirchner and Fouque were presented in Eurocrypt 2017 [KF17]. The first one is a new subfield attack which uses the subring structure and projection technique, applicable to power-of-two cyclotomic modulus and has similar result compared to previous two subfield attacks. On the other hand, the second analysis is about the lattice reduction algorithm on the full NTRU lattice with general modulus. Kirchner and Fouque demonstrated that the lattice reduction on the full NTRU lattice yield similar performance as in subfield attacks. The comparative analysis between this and our result is not yet clear.

Our Contribution. We suggest a new lattice algorithm to solve the overstretched NTRU problem. Our algorithm has a similar asymptotic time complexity to the subfields [ABD16,CJL16,KF17] when an expansion factor of the modulus polynomial $F(X)$ is $\mathrm{poly}(n)$.[1] To be specific, our algorithm solves the $\mathsf{NTRU}_{R,q,M,\frac{q}{2}}$ problem in time $\mathrm{poly}(n) \cdot 2^{\Theta(\beta)}$ by using BKZ of block size

[1] The expansion factor is defined as $\sup\limits_{\boldsymbol{a},\boldsymbol{b}\in K^*} \frac{\|\boldsymbol{a}\cdot\boldsymbol{b}\|}{\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|}$. Note that most of cryptographic schemes [Gen09,GGH13,BLLN13,LATV12] are set to be $C_F = \mathrm{poly}(n)$.

$\frac{\beta}{\log \beta} \sim \frac{27n \log M}{\log^2 q}$ which has only different in the constant term with respect to $\frac{\beta}{\log \beta} \sim \frac{16n \log M}{\log^2 q}$ of the subfield attacks. Our new algorithm indicates that the cryptanalyses of GGH [CJL16], LTV and YASHE [ABD16] schemes are still effective regardless of choosing a polynomial modulus of prime degree, opposed to the expectation in [BCLvV16].

In addition, we suggest an generalization of the subfield attacks suggested in [ABD16,CJL16]. Our attack reduces the required block sizes of the BKZ algorithm from β to β' satisfying $\frac{\beta'}{\log \beta'} \sim \frac{27n \log M}{2 \log^2 q}$, while $\frac{\beta}{\log \beta} \sim \frac{16n \log M}{\log^2 q}$ in the previous. Our second attack implies that the degree should be increased from n to $\frac{32n}{27}$ in order to maintain the same security level in the cryptosystems based on the overstretched NTRU problems.

Technical overview. An NTRU problem with an instance $\boldsymbol{h}(X) = [\boldsymbol{f}/\boldsymbol{g}]_q \in R_q := \mathbb{Z}[X]/\langle F(X), q\rangle$ for a polynomial $F(X)$ of degree n is associated with an NTRU lattice having the following basis matrix of Hermite normal form:

$$\boldsymbol{B} = \begin{pmatrix} q \cdot \boldsymbol{I}_n & \phi(\boldsymbol{h}) \\ \boldsymbol{O} & \boldsymbol{I}_n \end{pmatrix} \in \mathbb{Z}^{2n \times 2n},$$

where $\phi(\boldsymbol{h})$ is an n-by-n matrix whose (i,j) entry is X^j's coefficient of $\boldsymbol{h}(X) \cdot X^i$. By identifying a polynomial of degree $< n$ in R as a coefficient vector in $\mathbb{Z}^n$, we can see that this lattice contains $(\boldsymbol{f}, \boldsymbol{g})^T$ and a short vector in this lattice gives a solution of the NTRU problem. The lattice reduction algorithm to find a short vector in the NTRU lattice heavily relies on the dimension of the matrix. Therefore, the subfield attacks of recently suggested papers [CJL16,ABD16,KF17] focus on finding a relevant lattice of smaller dimension.

Our idea is based on the property of NTRU lattices having extremely short vectors $(\boldsymbol{f}, \boldsymbol{g})^T, (\boldsymbol{f} \cdot X, \boldsymbol{g} \cdot X)^T, \cdots, (\boldsymbol{f} \cdot X^{n-1}, \boldsymbol{g} \cdot X^{n-1})^T$. First, we consider the projection technique introduced by Fouque and Kirchner [KF17] that apply a natural projection ψ from $\mathbb{Z}^n$ to $\mathbb{Z}^s$ $(s < n)$ to a NTRU lattice L to obtain a lattice L'. By a lattice reduction algorithm, those short vectors can be found under the assumption that if a NTRU lattice L contains extremely short vectors, short vectors of projected lattice L' smaller than the Gaussian heuristic are projection of short vector of L. Then one can recover the original short vectors in the NTRU lattice by finding those vectors.

Fouque and Kirchner applied this technique to a subring lattice of the original NTRU lattice to solve the NTRU problem more efficiently. If a subring is of a sufficiently small dimension and still contains extremely short vectors, the projection technique combined with BKZ finds a short element from the subring. However, if a base modulus ring has no suitable subrings, applying the projection technique alone cannot achieve the asymptotic result of combined technique.

In this paper, instead, we locate and utilize a sublattice of an NTRU lattice containing extremely short vectors without using its ring structure. We show that the sublattice consisting of the first $(n+s)$ columns of the NTRU lattice L has a vector of size $\lambda_1(L)^{n/s} \approx \sqrt{\|\boldsymbol{f}\|^2 + \|\boldsymbol{g}\|^2}^{n/s}$. By applying the projection technique to this lattice, we obtain sufficiently short vector of the NTRU lattice in the same complexity of the previous subfield or subring attacks regardless of the ring structure.

The proposed algorithm has the similar form of asymptotic complexity poly$(n)\cdot 2^{\Theta(\beta)}$ [ABD16,CJL16], but with slightly larger leading coefficient in the exponent. Our second algorithm improves this in the case of power of 2 cyclotomic fields. We observe that the subfield algorithm can be applied similarly to $\boldsymbol{h}\cdot X^{-i}$ for any positive integer $i<n$.

The subfield attacks first find a pair $(\mathrm{Tr}(\boldsymbol{f}\cdot \mathrm{N}(\boldsymbol{g})/\boldsymbol{g}),\ \mathrm{N}(\boldsymbol{g}))$ by applying a lattice reduction to $\mathrm{Tr}(\boldsymbol{h}) = [\mathrm{Tr}(\boldsymbol{f}\cdot \mathrm{N}(\boldsymbol{g})/\boldsymbol{g})/\mathrm{N}(\boldsymbol{g})]_q$, where $\mathrm{N}(\cdot)$ is a norm function from $\mathbb{Q}[X]/\langle X^n+1\rangle$ to $\mathbb{Q}[X^m]/\langle X^n+1\rangle$ for a divisor m of n, and a $\mathrm{Tr}(\cdot)$ is a trace function. In our second algorithm, instead of applying lattice reduction to each NTRU lattice for $\boldsymbol{h}\cdot X^{-i}$ with basis matrix

$$\boldsymbol{B}' = \begin{pmatrix} q\cdot \boldsymbol{I}_m & \phi(\mathrm{Tr}(\boldsymbol{h})) \\ \boldsymbol{O} & \boldsymbol{I}_m \end{pmatrix} \in \mathbb{Z}^{2m\times 2m},$$

we consider the combined lattice with basis matrix:

$$\boldsymbol{B}'' = \begin{pmatrix} q\cdot \boldsymbol{I}_m & \phi(\mathrm{Tr}(\boldsymbol{h})) & \dots & \phi(\mathrm{Tr}(\boldsymbol{h}\cdot X^{-(\frac{n}{m}-1)})) \\ \boldsymbol{O} & \boldsymbol{I}_m & \dots & \boldsymbol{O} \\ \vdots & \vdots & \dots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \dots & \boldsymbol{I}_m \end{pmatrix} \in \mathbb{Z}^{(m+n)\times(m+n)}.$$

A short element of this lattice corresponds to a multiple of $\boldsymbol{g}$. This lattice has the same volume with, but a larger dimension than that of individual NTRU lattice. By taking a proper sublattice generated by the first r columns of $\boldsymbol{B}''$ $(m<r<m+n)$, we can reduce the reduction time to obtain an output.

Organization. In Section 2, we introduce some notation and preliminary information related to NTRU problems. In Section 3, we explain a new algorithm to solve the NTRU problems with a small expansion factor. In Section 4, we describe an improvement of the subfield attack.

2 Preliminaries

Notation. Throughout this paper, we cryptanalyze the NTRU problem over the ring $R=\mathbb{Z}[X]/\langle F(X)\rangle$. Here n is a degree of integral polynomial $F(X)$, and q is a positive integer. Moreover we define $K:=\mathbb{Q}[X]/\langle F(X)\rangle$. We use bold letters to denote vectors or ring elements in $\mathbb{Z}^n$ or R.

If m is a integer, we use the notations $m \bmod q$ or $[m]_q$ as an element in $\mathbb{Z}_q$, which congruent to m modulo q. On the contrary, we define $\iota:\mathbb{Z}_q\to\mathbb{Z}$ by $[m]_q\mapsto m$ for $-\frac{q}{2}<m\le\frac{q}{2}$. In general, we denote $[R]_q := R/qR = \mathbb{Z}_q[X]/\langle F(X)\rangle$.

For $\boldsymbol{u}=\sum_{i=0}^{n-1} u_i\cdot X^i \in K$ and $[\boldsymbol{u}]_q=\sum_{i=0}^{n-1}[u_i]_q\cdot X^i\in[R]_q$, we also denote $[\boldsymbol{u}]_q=\sum_{i=0}^{n-1}[u_i]_q\cdot X^i\in[R]_q$ and $\iota([\boldsymbol{u}]_q)=\sum_{i=0}^{n-1}\iota([u_i]_q)\cdot X^i\in R$.

If two functions of f, g grow asymptotically equal, or $f(x)/g(x) \to 1$ as $x \to \infty$, we denote $f \sim g$.

2.1 Lattices

A lattice $\mathcal{L} \subset \mathbb{R}^n$ is the set of all $\mathbb{Z}$-linear combinations of m linearly independent vectors $\boldsymbol{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_m\}$. We use $\mathcal{L}(\boldsymbol{B})$ to denote the lattice generated by the set $\boldsymbol{B}$. If $\boldsymbol{B}$ is a matrix, $\mathcal{L}(\boldsymbol{B})$ is the lattice generated by the set of columns of $\boldsymbol{B}$, and $\boldsymbol{B}$ is called a basis matrix of $\mathcal{L}(\boldsymbol{B})$. When $m = n$, this lattice is called a full-rank lattice. The determinant $\det(\mathcal{L})$ of lattice $\mathcal{L}$ is defined as $\det(\boldsymbol{B})$ for the basis matrix $\boldsymbol{B}$ of $\mathcal{L}$.

The principal ideal $\langle \boldsymbol{u} \rangle$ for $\boldsymbol{u} \in R$ corresponds to a lattice, which is deemed an ideal lattice, when $\boldsymbol{u} = \sum_{i=0}^{n-1} u_i \cdot X^i \in R$ with a column vector $(u_0, \cdots, u_{n-1})^T \in \mathbb{Z}^n$ is identified. $\{\boldsymbol{u}, \boldsymbol{u} \cdot X, \cdots, \boldsymbol{u} \cdot X^{n-1}\}$ is a basis of the ideal lattice of $\langle \boldsymbol{u} \rangle$.

Now we define a map $\phi : R \to \mathbb{Z}^{n \times n}$ by $\boldsymbol{u} \mapsto \left[\boldsymbol{u}, \boldsymbol{u} \cdot X, \cdots, \boldsymbol{u} \cdot X^{n-1}\right]$. Then $\phi(\boldsymbol{u})$ is a basis matrix of an ideal lattice $\langle \boldsymbol{u} \rangle$.

Successive minima For an m-rank lattice $\mathcal{L}$ and its shortest vector $\boldsymbol{c}_1$, we denote $\lambda_1(\mathcal{L}) := \|\boldsymbol{c}_1\|$ as the first successive minima of $\mathcal{L}$. Moreover, we inductively define successive minimas as follows: for a positive integer $k < m$ and given successive minimas $\lambda_1(\mathcal{L}), \cdots, \lambda_k(\mathcal{L})$ and correspond independent vectors $\boldsymbol{c}_1, \cdots, \boldsymbol{c}_k$, let $\boldsymbol{c}_{k+1}$ be a nonzero shortest vector in $\mathcal{L}$ independent to every $\boldsymbol{c}_i$ for $1 \leq i \leq k$. We also define the $(k+1)$-th successive minima $\lambda_{k+1}(\mathcal{L})$ as $\|\boldsymbol{c}_{k+1}\|$. Now let us present useful results about successive minimas.

Theorem (Minkowski). *Let $\mathcal{L}$ be an m-rank lattice. Then we have*

$$\lambda_1(\mathcal{L}) \leq \sqrt{m}\det(\mathcal{L})^{1/m}.$$

Heuristic (Gaussian Heuristic). *The size of successive minimas of an m-rank lattice $\mathcal{L}$ asymptotically as follows*

$$\forall 1 \leq i \leq m, \lambda_i(\mathcal{L}) \sim \sqrt{\frac{m}{2\pi e}}\det(\mathcal{L})^{1/m}.$$

[Ajt06] showed that a random lattice satisfies the above equation with overwhelming probability.

Lattice reduction algorithm To find a short element of a lattice, lattice reduction algorithms such as the LLL algorithm and the BKZ algorithm are described in [LLL82,HPS11,ADRSD15]. These algorithms lead us to find an approximately short vector of a lattice with bounded time.

In the case of the BKZ algorithm, by [HPS11], the quality of the algorithms relies on the block size β. More precisely, using the BKZ algorithm upon basis $\boldsymbol{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n\}$, we obtain a reduced basis $\boldsymbol{B}' = \{\boldsymbol{b}'_1, \boldsymbol{b}'_2, \cdots, \boldsymbol{b}'_n\}$, which satisfies:

- $\|\boldsymbol{b}_1'\| \leq 2(\gamma_\beta)^{\frac{n-1}{2(\beta-1)}+\frac{3}{2}} \cdot (\det \mathcal{L})^{\frac{1}{n}}$ in poly$(n, size(\boldsymbol{B})) \cdot \mathcal{C}_{HKZ}(\beta)$ times or
- $\|\boldsymbol{b}_1'\| \leq 4(\gamma_\beta)^{\frac{n-1}{\beta-1}+3} \cdot (\varLambda_1(\mathcal{L}))^{\frac{1}{n}}$ in poly$(n, size(\boldsymbol{B})) \cdot \mathcal{C}_{HKZ}(\beta)$ times,

where $\mathcal{L}$ is the lattice $\mathcal{L}(\boldsymbol{B})$, $\gamma_\beta \leq \beta$ is the Hermite constant of rank β, $size(\boldsymbol{B})$ is the size of the largest entries of the basis matrix $\boldsymbol{B}$, and $\mathcal{C}_{HKZ}(\beta) = 2^{O(\beta)}$ is the cost of HKZ-reduction in dimension β. For an output $\boldsymbol{b}_1'$ of lattice reduction algorithms, we call $\dfrac{\boldsymbol{b}_1'}{\det(\mathcal{L})^{1/n}}$ and $\dfrac{\boldsymbol{b}_1'}{\lambda_1(\mathcal{L})}$ as a Hermite factor and approximate factor, respectively.

For convenience of calculation, throughout this paper, we assume that we have a lattice reduction algorithm $\mathcal{A}_\delta$, whose output contains a short vector $\boldsymbol{b}$ with Euclidean norm less than $\delta^n \cdot \det(\mathcal{L})^{1/n}$ or $\delta^{2n} \cdot \lambda_1(\mathcal{L})$ for an n-dimensional lattice $\mathcal{L}$ instead of $2(\gamma_\beta)^{\frac{n-1}{2(\beta-1)}+\frac{3}{2}} \cdot (\det \mathcal{L})^{\frac{1}{n}}$ or $4(\gamma_\beta)^{\frac{n-1}{\beta-1}+3} \cdot (\varLambda_1(\mathcal{L}))^{\frac{1}{n}}$, respectively.

2.2 Euclidean norm of $\boldsymbol{K}$

For $\boldsymbol{u} = \sum_{i=0}^{n-1} u_i \cdot X^i \in \mathbb{Z}[X]/\langle F(X)\rangle$, we define $\|\boldsymbol{u}\|$ is the Euclidean norm of the vector $(u_0, \cdots, u_{n-1})$, and denote $\|[\boldsymbol{u}]_q\| = \|\iota(\boldsymbol{u}_q)\|$. For inverses of $\boldsymbol{g}$ over $[R]_q$ and K, in general, $\|[\boldsymbol{g}]_q^{-1}\| \neq \|\boldsymbol{g}^{-1}\|$. Thus, we use the notation $\|\boldsymbol{g}^{-1}\|_K$ instead of $\|\boldsymbol{g}^{-1}\|$ to avoid confusions.

Next, we show that the bound of $\|\boldsymbol{a}+\boldsymbol{b}\|$ and $\|\boldsymbol{a}\cdot\boldsymbol{b} \bmod F(X)\|$ for $\boldsymbol{a}, \boldsymbol{b} \in K$. The following triangle inequality holds:

Lemma 1. *$\|\boldsymbol{a}+\boldsymbol{b}\| \leq \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$ for any $\boldsymbol{a}, \boldsymbol{b} \in K$.*

We present a definition related to the multiplication.

Definition. For a polynomial $F(X)$ and ring $K = \mathbb{Q}[X]/\langle F(X)\rangle$, the expansion factor of F is defined by

$$C_F := \sup_{\boldsymbol{a},\boldsymbol{b} \in K^*} \frac{\|\boldsymbol{a}\cdot\boldsymbol{b}\|}{\|\boldsymbol{a}\| \cdot \|\boldsymbol{b}\|}.$$

Then clearly inequality $\|\boldsymbol{a}\cdot\boldsymbol{b}\| \leq \|\boldsymbol{a}\| \cdot \|\boldsymbol{b}\| \cdot C_F$ holds. Note that C_F is bounded as polynomial of the degree n of F in general, which is induced by following lemma.

Lemma 2 ([Gen09], Theorem 9). *Let $K = \mathbb{Q}[X]/\langle F(X)\rangle$ and suppose $F(X) = X^n + h(X)$ where $h(X)$ has degree at most $n-(n-1)/k$ for $k \geq 2$. Then,*

$$C_F \leq \sqrt{2n} \cdot \left(1 + 2n \cdot (\sqrt{(k-1)n} \cdot \|F(X)\|)^k\right).$$

In the case of $F(X) = X^n \pm 1$ or $X^n - X - 1$, C_F is bounded by $\sqrt{n}$ or $2\sqrt{n}$, respectively.

2.3 NTRU problem and its lattice based approach

In this section, we describe the NTRU problem and its related lemmas. First of all, we state an NTRU problem as follows:

Problem 1 (The NTRU problem)
Let n and q be integers, M be a positive real number and $F(X)$ be a degree n integral polynomial. For a polynomial ring $R := \mathbb{Z}[X]/\langle F(X)\rangle$, $\boldsymbol{f}$ and $\boldsymbol{g}$ are sampled from R and have Euclidean norms bounded by M. For given a polynomial $\boldsymbol{h} = [\boldsymbol{f}/\boldsymbol{g}]_q$, the NTRU problem $\mathsf{NTRU}_{R,q,M,\tau}$ is to find $\boldsymbol{a}\cdot\boldsymbol{f}$, $\boldsymbol{a}\cdot\boldsymbol{g} \in R$ for some $\boldsymbol{a} \in R$, such that $\|\boldsymbol{a}\cdot\boldsymbol{f}\|, \|\boldsymbol{a}\cdot\boldsymbol{g}\| \le \tau$.

In many NTRU-based applications, M is taken to be similar to poly(n). Furthermore, when q is set to be super-polynomial in n, the NTRU problem is called the *overstretched* NTRU problem. In [ABD16,CJL16,KF17], the authors suggested subfield attack to solve overstretched NTRU problem on $R = \mathbb{Z}[X]/\langle X^n + 1\rangle$ for a power of two n with $\tau = q$. In this paper we focus on the overstretched NTRU problem on $R = \mathbb{Z}[X]/\langle F(X)\rangle$ for $C_F = \text{poly}(n)$ with $\tau = q/2$ and $M = \text{poly}(n)$. Next, we state an useful lemma to solve the NTRU problem.

Lemma 3 ([GGH13], Lemma 3). *Let $\boldsymbol{f}, \boldsymbol{g} \in R = \mathbb{Z}[X]/\langle F(X)\rangle$ be relative prime and $[\boldsymbol{g}]_q$ is invertible in $[R]_q = R/qR$. If $\boldsymbol{c} \in R$ satisfies $\|\boldsymbol{c}\| < \dfrac{q}{2C_F \cdot \|\boldsymbol{f}\|}$ and $\|[\boldsymbol{c}\cdot\boldsymbol{f}\cdot\boldsymbol{g}^{-1}]_q\| < \dfrac{q}{2C_F\cdot\|\boldsymbol{g}\|}$, then $\boldsymbol{c}$ and $[\boldsymbol{c}\cdot\boldsymbol{f}\cdot\boldsymbol{g}^{-1}]_q$ are contained in the ideal $\langle\boldsymbol{g}\rangle$ and $\langle\boldsymbol{f}\rangle$, respectively.*

Proof. Let $\boldsymbol{w} := [\boldsymbol{c}\cdot\boldsymbol{f}\cdot\boldsymbol{g}^{-1}]_q$. Then, $[\boldsymbol{g}\boldsymbol{w}]_q = [\boldsymbol{c}\boldsymbol{f}]_q$. Since $\|\boldsymbol{w}\| < q/(2\|\boldsymbol{g}\|\cdot C_F)$, we have $\|\boldsymbol{g}\boldsymbol{w}\| \le \|\boldsymbol{g}\|\cdot\|\boldsymbol{w}\|\cdot C_F \le q/2$ and $\|\boldsymbol{c}\boldsymbol{f}\| \le \|\boldsymbol{c}\|\cdot\|\boldsymbol{f}\|\cdot C_F \le q/2$. Therefore, $\boldsymbol{g}\boldsymbol{w} = \boldsymbol{c}\boldsymbol{f}$ in $\mathbb{Z}[X]/\langle F(X)\rangle$. Because $\boldsymbol{c}\boldsymbol{f} \in \langle\boldsymbol{g}\rangle$ and $\boldsymbol{f}$ is a relative prime to $\boldsymbol{g}$, we can conclude $\boldsymbol{c} \in \langle\boldsymbol{g}\rangle$. With similar reasons, we have $\boldsymbol{w} \in \langle\boldsymbol{f}\rangle$. □

Hence, **Lemma 3** gives if one can find a short pair $([\boldsymbol{c}\cdot\boldsymbol{h}]_q, \boldsymbol{c})$ with an Euclidean norm smaller than $\dfrac{q}{2\cdot C_F\cdot M}$, one will be able to solve the $\mathsf{NTRU}_{R,q,M,\frac{q}{2}}$.

By employing the above property, we can describe the basic lattice-based approach to solve the NTRU problem with an input polynomial $\boldsymbol{h} = [\boldsymbol{f}/\boldsymbol{g}]_q$. Let consider the lattice $\mathcal{L}$ generated by following $2n \times 2n$ basis matrix:

$$\boldsymbol{B} = \begin{pmatrix} q\cdot\boldsymbol{I}_n & \phi(\boldsymbol{h}) \\ \boldsymbol{O} & \boldsymbol{I}_n \end{pmatrix} \in \mathbb{Z}^{2n\times 2n}.$$

For a polynomial $\boldsymbol{c} = \sum_{i=0}^{n-1} c_i \cdot X^i$, the polynomial vector $(\boldsymbol{c}\cdot\boldsymbol{h}, \boldsymbol{c}) = \sum_{i=0}^{n-1} c_i\cdot(X^i\cdot\boldsymbol{h}, X^i)$ corresponds to a lattice point $\sum_{i=0}^{n-1} c_i \cdot B_{n+i}$, where B_{n+i} is the $n+i$-th column vector of the basis matrix $\boldsymbol{B}$. It implies that $([\boldsymbol{c}\cdot\boldsymbol{h}]_q, \boldsymbol{c})$ is also identified to lattice point of $\mathcal{L}(\boldsymbol{B})$. Hence, finding a short pair $([\boldsymbol{c}\cdot\boldsymbol{h}]_q, \boldsymbol{c})$ is the same as finding a short lattice point of $\mathcal{L}(\boldsymbol{B})$.

We also state another lemma to solve NTRU problems. This is applicable to the pair $(\boldsymbol{a}, \boldsymbol{b})$ where $\boldsymbol{b}$ is known to be a multiple of $\boldsymbol{g}$.

Lemma 4. *Let $\boldsymbol{g}$ be an element of $R = \mathbb{Z}[X]/\langle F(X)\rangle$ and $\boldsymbol{f} \in R$ be relative prime to $\boldsymbol{g}$. For some $\boldsymbol{d} \in R$, if $\boldsymbol{d} \cdot \boldsymbol{g} \in \langle \boldsymbol{g} \rangle \subset R$ satisfies $\|\boldsymbol{d} \cdot \boldsymbol{g}\| < \dfrac{q}{2 \cdot C_F^2 \cdot \|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\|}$ then $\boldsymbol{d} \cdot \boldsymbol{g}$ and $[\boldsymbol{d} \cdot \boldsymbol{g} \cdot \boldsymbol{h}]_q$ are solution of $\mathsf{NTRU}_{R,q,M,\frac{q}{2}}$ problem with input $\boldsymbol{h}$.*

Proof. By conditions, we have

$$\|\boldsymbol{d} \cdot \boldsymbol{f}\| = \|\boldsymbol{d} \cdot \boldsymbol{f} \cdot \boldsymbol{g}^{-1} \cdot \boldsymbol{g}\| \leq C_F^2 \|\boldsymbol{d} \cdot \boldsymbol{g}\| \cdot \|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\| < q/2.$$

Hence, $[\boldsymbol{d} \cdot \boldsymbol{g} \cdot \boldsymbol{h}]_q$ has the form $\boldsymbol{d} \cdot \boldsymbol{f}$. □

3 A general attack on the NTRU problem

We first present the result of our general attack as follows:

Theorem 1 (Heuristic) *Let $\boldsymbol{f}$ and $\boldsymbol{g}$ be relatively prime elements of $R = \mathbb{Z}[X]/\langle F(X)\rangle$. If q satisfies* [2]

$$2\sqrt{\frac{n}{\pi e}} \cdot \|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\|_K \cdot C_F^2 \leq q^{1/3},$$

one can solve the $\mathsf{NTRU}_{R,q,M,q/2}$ problems upon $[\boldsymbol{f}/\boldsymbol{g}]_q$ in $\mathrm{poly}(n) \cdot 2^{O(\beta)}$ time by using the BKZ algorithm with block size β with $\beta/\log\beta \geq \frac{27n\log M}{\log^2 q} + o(1)$.

In the case of the overstretched $\mathsf{NTRU}_{R,q,M,q/2}$ problems, the inequality condition is asymptotically satisfied if the value of $\|\boldsymbol{g}^{-1}\|_K, M$ and C_F are $\mathrm{poly}(n)$ size.

Our main strategy for solving the NTRU problem is to use the projection technique on a sublattice of the NTRU lattice. Note that the projection technique is effective in a lattice having two properties: first, it has a very short vector so that its projection is still far smaller than the other short vectors implied by the Gaussian Heuristic. Secondly, The short vector in the original lattice can be computed efficiently from the output vectors of the projection technique. The first is assumed by the following statement:

Assumption ([KF17]). *Let $\mathcal{L}$ be an n-dimensional lattice contains extremely short vectors compared to $\sqrt{\frac{n}{2\pi e}} \cdot \det(\mathcal{L})^{1/n}$, the expected size of successive minimas by the Gaussian Heuristic, and $\psi : \mathbb{Z}^n \to \mathbb{Z}^m$ a projection map. If the vector $\boldsymbol{b} \in \psi(\mathcal{L})$ is shorter than $\sqrt{\frac{m}{2\pi e}} \cdot \det(\psi(\mathcal{L}))^{1/m}$, then $\boldsymbol{b}$ is a projection of an extremely short vector of $\mathcal{L}$.*

The above assumption experimentally shown to hold in the NTRU lattice by [KF17]. The second is easily conducted in the NTRU lattice because of its algebraic structure.

[2] We can change the condition of q into $2\|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\|_K \cdot C_F^2 \leq q^{1-\frac{2}{3\sqrt{6}}}$. Refer to the **Appendix A** for its details.

Now the existence of short vectors is needed to use the **Assumption**. The authors of [KF17] prove the existence in the lattice induced from subring by using the algebraic structure, and exploit the projection technique on that lattice to solve the NTRU problem. On the other hand, we show that some sublattice of the NTRU lattice has a vector which has the similar size to the previous one, and use the projection technique on the sublattice.

3.1 The existence of the short vector in the sublattice

By showing the existence of the short vector in the sublattice of the NTRU lattice, we can give the bound the λ_1 of the projection of the sublattice.

Now we present the lemma inspired by [CL15] to show that the sublattice of the NTRU lattice has a sufficiently short vector.

Lemma 5. *Let* $\boldsymbol{h} = [\boldsymbol{f}/\boldsymbol{g}]_q \in [R]_q$, $\boldsymbol{B} = \begin{bmatrix} q \cdot \boldsymbol{I}_n & \phi(\boldsymbol{h}) \\ 0 & \boldsymbol{I}_n \end{bmatrix}$ *is a* $2n \times 2n$ NTRU *matrix and* $\boldsymbol{B}_{n+s}$ *the* $2n \times (n+s)$ *submatrix of* $\boldsymbol{B}$ *obtained by removing the last* $(n-s)$ *columns from* $\boldsymbol{B}$*. Then* $\mathcal{L}(\boldsymbol{B}_{n+s})$ *is a sublattice of* $\mathcal{L}(\boldsymbol{B})$ *and contains a vector of length* $\le \sqrt{s} \cdot (\prod_{i=1}^{n} \|(\boldsymbol{f} \cdot X^i, \boldsymbol{g} \cdot X^i)\|)^{1/s}$.

Proof. Let U be an unimodular matrix such that $\phi(\boldsymbol{g}) \cdot U = HNF(\phi(\boldsymbol{g}))$ where $HNF(\phi(\boldsymbol{g}))$ is an upper triangular Hermite normal form of $\phi(\boldsymbol{g})$. Then, we have

$$\boldsymbol{L} = \mathcal{L}\left(\begin{bmatrix} \phi(\boldsymbol{f}) \\ \phi(\boldsymbol{g}) \end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix} \phi(\boldsymbol{f}) \cdot U \\ \phi(\boldsymbol{g}) \cdot U \end{bmatrix}\right) \subset \mathcal{L}(\boldsymbol{B}),$$

and the determinant of the lattice $\boldsymbol{L}$ is bounded by $\prod_{i=1}^{n} \|(\boldsymbol{f} \cdot X^i, \boldsymbol{g} \cdot X^i)\|$. Let $\boldsymbol{L}'$ be a sublattice of $\boldsymbol{L}$, which is generated by first s columns of $\boldsymbol{L} = \mathcal{L}\left(\begin{bmatrix} \phi(\boldsymbol{f}) \cdot U \\ \phi(\boldsymbol{g}) \cdot U \end{bmatrix}\right)$. Then, determinant of $\boldsymbol{L}'$ is smaller than that of $\boldsymbol{L}$. According to Minkowski's theorem on its sublattice $\mathcal{L}'$, we get $\lambda_1(\mathcal{L}') \le \sqrt{s} \cdot \det(\mathcal{L}')^{1/s} \le \sqrt{s} \cdot (\prod_{i=1}^{n} \|(\boldsymbol{f} \cdot X^i, \boldsymbol{g} \cdot X^i)\|)^{1/s}$.

Now, let $(\boldsymbol{c} \cdot \boldsymbol{f}, \boldsymbol{c} \cdot \boldsymbol{g})^T$ be the shortest vector of $\boldsymbol{L}' \subset \boldsymbol{L}$. Since the i-th coefficients of $(\boldsymbol{c} \cdot \boldsymbol{f}, \boldsymbol{c} \cdot \boldsymbol{g})$, for $n+s+1 \le i \le 2n$, are zero, it is also an element of $\boldsymbol{B}_{n+s}$. Hence, we can obtain $\lambda_1(\mathcal{L}(\boldsymbol{B}_{n+s})) \le \|(\boldsymbol{c} \cdot \boldsymbol{f}, \boldsymbol{c} \cdot \boldsymbol{g})\| \le \sqrt{s} \cdot (\prod_{i=1}^{n} \|(\boldsymbol{f} \cdot X^i, \boldsymbol{g} \cdot X^i)\|)^{1/s}$. □

3.2 Projection technique on $\mathcal{L}(\boldsymbol{B}_{n+s})$

Now we can use the projection technique on the sublattice by virtue of the existence of short vector by **Lemma 5**. We consider the following matrix:

$$\widetilde{\boldsymbol{B}}_{m,s} = \begin{pmatrix} q \cdot \boldsymbol{I}_m & \psi_{m,s}(\phi(\boldsymbol{h})) \\ 0 & \boldsymbol{I}_s \end{pmatrix}$$

where the projection map $\psi_m : R \to \mathbb{Z}^m$ defined by $\boldsymbol{u} \mapsto (u_0, u_1, \cdots, u_{m-1})^T$ and its extension $\psi_{m,s}(\phi(\boldsymbol{u})) = [\psi_m(\boldsymbol{u}), \psi_m(\boldsymbol{u} \cdot X), \cdots, \psi_m(\boldsymbol{u} \cdot X^{s-1})]$.

Then $\mathcal{L}(\widetilde{\boldsymbol{B}}_{m,s})$ is the projected lattice of $\mathcal{L}(\boldsymbol{B}_{n+s})$ with respect to the last $(m+s)$ coordinates. To clarify the descriptions, we state the adapted **Assumption** of the form we need.

Assumption*. *If a vector $\boldsymbol{b} \in \mathcal{L}(\widetilde{\boldsymbol{B}}_{m,s})$ is shorter than $\sqrt{\frac{m+s}{2\pi e}} \cdot q^{\frac{m}{m+s}}$ which is the expected size of successive minimas of $\mathcal{L}(\widetilde{\boldsymbol{B}}_{m,s})$ by the Gaussian Heuristic, then $\boldsymbol{b}$ must be a projection of short multiple of $(\boldsymbol{f}, \boldsymbol{g})^T$.*

By the **Assumption***, we find a short vector in $\mathcal{L}(\widetilde{\boldsymbol{B}}_{m,s})$ to solve the NTRU problem. The lattice reduction algorithm $\mathcal{A}_\delta$ upon $\widetilde{\boldsymbol{B}}_{m,s}$ outputs a vector $\boldsymbol{b}$ such that

$$\|\boldsymbol{b}\| \le \delta^{2m+2s} \cdot \lambda_1(\widetilde{\boldsymbol{B}}_{m,s}) \le \delta^{2m+2s} \cdot \sqrt{s} \cdot (\|(\boldsymbol{f}, \boldsymbol{g})\|)^{n/s}.$$

The second inequality comes from **Lemma 5**. If this size is smaller than $\sqrt{\frac{m+s}{2\pi e}} \cdot q^{\frac{m}{m+s}}$, which is an expected size of successive minimas by the Gaussian heuristic, $\boldsymbol{b}$ would be of the form $(\psi_m(\boldsymbol{d} \cdot \boldsymbol{f}), \boldsymbol{d} \cdot \boldsymbol{g})^T$ for some polynomial $\boldsymbol{d}$ by **Assumption***. Therefore, one can recover $\boldsymbol{d} \cdot \boldsymbol{g}$. If the size of $\boldsymbol{d} \cdot \boldsymbol{g}$ is smaller than $\dfrac{q}{2\|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\|_K \cdot C_F^2}$, the $\mathsf{NTRU}_{R,q,M,q/2}$ problem can be solved by **Lemma 4**.

This process can be randomized using the basis order change technique used in [CL15].

Optimizing condition of the projection technique

In this section, we explain how to choose the best dimension of sublattice and projection to solve the NTRU problem by investigating the condition of **Lemma 4**.

Note that the inequality

$$\delta^{2m+2s} \cdot \sqrt{s} \cdot (\sqrt{2}M)^{n/s} \le \sqrt{\frac{m+s}{2\pi e}} \cdot q^{\frac{m}{m+s}} \tag{1}$$

implies that we can solve the $\mathsf{NTRU}_{R,q,M,q/2}$ problem by **Lemma 4**, because

$$\|\boldsymbol{d} \cdot \boldsymbol{g}\| \le \|\boldsymbol{b}\| \le \delta^{2m+2s} \cdot \sqrt{s} \cdot (\|(\boldsymbol{f}, \boldsymbol{g})\|)^{n/s} \le \delta^{2m+2s} \cdot \sqrt{s} \cdot (\sqrt{2}M)^{n/s}$$

holds. To elicit the optimizing condition, we observe the inequality (1).

After taking the log of (1), we get

$$2(m+s) \cdot \log \delta + \frac{\log s}{2} + \frac{n}{s} \cdot \log \sqrt{2}M \le \log \sqrt{\frac{m+s}{2\pi e}} + \frac{m}{m+s} \cdot \log q$$

or, by approximating $\log \sqrt{\frac{n+s}{2\pi e}} - \frac{\log s}{2} = 0$,

$$2(m+s) \cdot \log \delta + \frac{n}{s} \cdot \log \sqrt{2}M + \frac{s}{m+s} \cdot \log q \le \log q.$$

By adopting arithmetic and geometric mean inequality, we finally achieve the condition of δ as

$$54n \log \delta \log \sqrt{2}M \leq \log^2 q.$$

In order to further specify this, we need to choose m, s as the following equality conditions :

$$s = \sqrt[3]{\frac{n^2 \log^2\left(\sqrt{2}M\right)}{2 \log q \cdot \log \delta}},\ m + s = \sqrt[3]{\frac{n \log q \log\left(\sqrt{2}M\right)}{4 \log^2 \delta}}.$$

If we use the BKZ lattice reduction with β for $\beta/\log \beta \geq \frac{27n \log M}{\log^2 q} + o(1)$, which implies $\log \delta \leq \frac{\log^2 q}{54n \cdot \log \sqrt{2}M}$, we get $\frac{s}{m+s} \leq 1/3$. Then we get the multiple of $\boldsymbol{g}$ shorter than $\sqrt{\frac{m+s}{2\pi e}} \cdot q^{2/3} \leq \sqrt{\frac{n}{\pi e}} \cdot q^{2/3}$, and we can prove the **Theorem 1** by adopting **Lemma 4**.

4 Algorithm on NTRU problems using a subfield

In this section, we describe a new algorithm to solve the NTRU problems for a ring $R = \mathbb{Z}[X]/\langle X^n + 1\rangle$ with a power of two n upon $\boldsymbol{h} = [\boldsymbol{f}/\boldsymbol{g}]_q$. Therefore, the ring R is the ring of integer of $K = \mathbb{Q}[X]/\langle X^n + 1\rangle$. We denote $K_t = \mathbb{Q}[X^{2^t}]/\langle X^n+1\rangle$, $R_t = \mathbb{Z}[X^{2^t}]/\langle X^n+1\rangle$, and $n_t = n/2^t$. Then the trace $\mathrm{Tr}_{K/K_t}(\cdot)$ and norm map $\mathrm{N}_{K/K_t}(\cdot)$ are well defined.

The subfield attack [ABD16,CJL16] uses only one polynomial $\mathrm{N}_{K/K_t}(\boldsymbol{h})$ or $\mathrm{Tr}_{K/K_t}(\boldsymbol{h})$. Instead, we use several polynomials rather than a single one. Then we have the following theorem.

Theorem 2 *Let n be a power of two and $\boldsymbol{g}$ an element of $R = \mathbb{Z}[X]/\langle X^n + 1\rangle$ with square free algebraic norm and $\boldsymbol{f} \in R$ a relatively prime to $\boldsymbol{g}$. When the sizes of M and $\|\boldsymbol{g}^{-1}\|_K$ are $\mathrm{poly}(n)$, and q is super-polynomial in n, one can solve the $\mathsf{NTRU}_{R,q,M,q/2}$ problems upon $[\boldsymbol{f}/\boldsymbol{g}]_q$ using the BKZ algorithm with block size β with $\beta/\log \beta \geq \frac{27n \log M}{2 \log^2 q} + o\left(\frac{n \log n \log M}{\log^3 q}\right)$ in $\mathrm{poly}(n) \cdot 2^{O(\beta)}$ time.*

By **Theorem 2**, the NTRU problem can be solved for a larger n than the subfield algorithm in same time complexity.

Generally, it is expected that it would be easier to recover $\boldsymbol{g}$, if several NTRU instances $\boldsymbol{h}_i = [\boldsymbol{f}_i/\boldsymbol{g}]_q$ are given instead of one element. Unfortunately, there is no algorithm that uses multiple instances to the best of our knowledge.

When an overstretched NTRU instance $[\boldsymbol{f}/\boldsymbol{g}]_q$ with square free norm $\mathrm{N}(\boldsymbol{g})$ is given, we provide a new algorithm that employs several polynomials in the form of $\mathrm{Tr}(\boldsymbol{h} \cdot X^{-i})$. More precisely, we prove that if there exist polynomials $\boldsymbol{c}_i \in R_t$ such that the size of $\{\boldsymbol{c}_i\}_{i=0}^{n_t-1}$ and $\sum_{i=0}^{n_t-1} \boldsymbol{c}_i \cdot \mathrm{Tr}(\boldsymbol{h} \cdot X^{-i})/2^t$ is small, $\sum_{i=0}^{n_t-1} \boldsymbol{c}_i \cdot X^{-i}$ is a multiple of $\boldsymbol{g}$.[3] The proof reduces the NTRU problem to finding a short vector

[3] In particular if $\boldsymbol{c}_i$ is equal to zero for $1 \leq i \leq n_t - 1$, it is an original subfield attack.

in an appropriate lattice. By choosing the optimized dimension, we can get the above theorem.

Now we start proving the **Theorem 2**. Let $\mu_{j,t}(\boldsymbol{a}) = \dfrac{\mathrm{Tr}_{K/K_t}(\boldsymbol{a} \cdot X^{-j})}{2^t}$ for given t. Then $\boldsymbol{a} = \sum_{i=0}^{n-1} a_i \cdot X^t \in R$ could be expressed as $\sum_{j=0}^{2^t-1} \mu_{j,t}(\boldsymbol{a}) \cdot X^j$. Using a μ notation, the $\boldsymbol{h}$ is of the form $\sum_{j=0}^{2^t-1} \mu_{j,t}(\boldsymbol{h}) \cdot X^j$. Our strategy is to use several polynomials $\mu_{0,t}(\boldsymbol{h}),\ \mu_{1,t}(\boldsymbol{h}),\ \cdots,\ \mu_{m-1,t}(\boldsymbol{h})$ rather than a single one. From now on, we use μ_j instead of $\mu_{j,t}$ for simplicity. Then, we can have the following lemma, extended from **Lemma 3**.

Lemma 6. *Let q be an integer and n be a power of two and $n_t = n/2^t$, $\boldsymbol{f}, \boldsymbol{g} \in R = \mathbb{Z}[X]/\langle X^n + 1\rangle$ and $\boldsymbol{g}$ has a square free norm $\mathrm{N}_{K/\mathbb{Q}}(\boldsymbol{g})$, $\boldsymbol{h} = [\boldsymbol{f}/\boldsymbol{g}]_q \in R_q$. If $\boldsymbol{c}_i \in R_t = \mathbb{Z}[X^{2^t}]/\langle X^n + 1\rangle$ for $0 \le i < 2^t$ satisfies the following inequalities:*

$$\|\boldsymbol{c}_i\| < \frac{q}{2^{t+1} \cdot C_F^3 \cdot \|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\|_K \cdot \|\boldsymbol{g}\|^{2^t}} \quad \textit{for all} \quad i$$

$$\left\| \left[\sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot \mu_i(\boldsymbol{h}) \right]_q \right\| < \frac{q}{2C_F \cdot \|\boldsymbol{g}\|^{2^t}},$$

then $\boldsymbol{c} = \sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot X^{-i}$ is contained in the ideal $\langle \boldsymbol{g} \rangle$, and the pair $(\boldsymbol{c},\ [\boldsymbol{c} \cdot \boldsymbol{f} \cdot \boldsymbol{g}^{-1}]_q)$ is a solution of $\mathsf{NTRU}_{R,q,M,q/2}$.

The proof of the **Lemma 6** is placed in **Appendix B**.

Next, we consider the following matrix for $\boldsymbol{h} = \sum_{i=0}^{2^t-1} \mu_i(\boldsymbol{h}) \cdot X^i$ so that we find such a vector $\boldsymbol{c} \in R$

$$\widehat{\boldsymbol{B}}_t = \begin{pmatrix} q \cdot \boldsymbol{I}_{n_t} & \phi_t(\mu_0(\boldsymbol{h})) & \phi_t(\mu_1(\boldsymbol{h})) & \cdots & \phi_t(\mu_{2^t-1}(\boldsymbol{h})) \\ 0 & \boldsymbol{I}_{n_t} & 0 & \cdots & 0 \\ 0 & 0 & \boldsymbol{I}_{n_t} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \boldsymbol{I}_{n_t} \end{pmatrix},$$

where $\phi_t(\mu_i(\boldsymbol{h}))$ is a basis matrix corresponding to the ideal lattice $\langle \mu_i(\boldsymbol{h}) \rangle$ over $\mathbb{Z}[X^{n_t}]/\langle X^n + 1\rangle$. Suppose that one can find a lattice point $\boldsymbol{b} = (\boldsymbol{b}' || \boldsymbol{b}_0 || \cdots || \boldsymbol{b}_{2^t-1}) \in \mathcal{L}(\widehat{\boldsymbol{B}}_t)$ such that $\|\boldsymbol{b}\| \le \dfrac{q}{2^{t+1} \cdot C_F^3 \cdot \|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\|_K \cdot \|\boldsymbol{g}\|^{2^t}}$. Then, trivially, $\|\boldsymbol{b}_i\| \le \|\boldsymbol{b}\|$ and $\left\| \left[\sum_{i=0}^{2^t-1} \boldsymbol{b}_i \cdot \mu_i(\boldsymbol{h}) \right]_q \right\| = \|\boldsymbol{b}'\| \le \|\boldsymbol{b}\|$. Hence, the short vector of $\mathcal{L}(\widehat{\boldsymbol{B}}_t)$ guarantee to find a vector satisfying the condition of **Lemma 6**.

To find a short lattice point, we apply the lattice reduction algorithms A_δ to a sublattice L' generated by the first $n_t + m$ column vector of the matrix $\widehat{\boldsymbol{B}}_t$. Therefore if we have:

$$\delta^{n_t+m} q^{\frac{n_t}{n_t+m}} \leq \frac{q}{2^{t+1} \cdot C_F^3 \cdot \|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\|_K \cdot \|\boldsymbol{g}\|^{2^t}},$$

we can find a lattice point $\boldsymbol{b}$ as we want.

Finally, in order to achieve the optimizing condition, one can get the following inequality by taking the logarithm function on both side. Therefore we have the following asymptotic inequality:

$$(n_t + m)\log\delta + \frac{n}{n_t}\log M + \frac{n_t}{n_t + m}\log q \leq \log q + o(\log n),$$

which is similar to **Section 3.2**. With similar method, $\log\delta \leq \frac{\log^2 q}{27n\log M}$ is the condition of δ to solve the NTRU problem. Hence, we can solve $\mathsf{NTRU}_{R,q,M,\frac{q}{2}}$ in $\mathrm{poly}(n) \cdot 2^{O(\beta)}$ time if $\beta/\log\beta \geq \frac{27n\log M}{2\log^2 q} + o\left(\frac{n\log n\log M}{\log^3 q}\right)$, using the BKZ algorithm with block size β.

5 Conclusion

In this work, we presented a new algorithm to solve the NTRU problem with small expansion factor. This algorithm is provided by applying the projection technique to a sublattice having an extremely short vector with an assumption. In the case of NTRU lattice, one can recover the solution from the output of the projection technique by using the algebraic structure. If a lattice has extremely short vectors and can process efficiently recovering, the projection technique would be a new approach to solve the problem related to that lattice.

In addition, we also suggested an improved algorithm to solve the NTRU problem faster than the previous algorithms when the base modulus is a cyclotomic polynomial with a smooth degree. The advantage of this algorithm comes from utilizing several polynomials. Unfortunately, this is not applicable to a general NTRU ring at the moment. It would be an interesting problem to investigate an algorithm to use multiple instances in the general case.

References

[ABD16] Martin Albrecht, Shi Bai, and Léo Ducas. A subfield lattice attack on overstretched ntru assumptions. In *Annual Cryptology Conference*, pages 153–178. Springer, 2016.

[ACLL14] Martin R Albrecht, Catalin Cocis, Fabien Laguillaumie, and Adeline Langlois. Implementing candidate graded encoding schemes from ideal lattices. In *Advances in Cryptology–ASIACRYPT 2015*, pages 752–775. Springer, 2014.

[ADRSD15] Divesh Aggarwal, Daniel Dadush, Oded Regev, and Noah Stephens-Davidowitz. Solving the shortest vector problem in 2 n time using discrete gaussian sampling. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, pages 733–742. ACM, 2015.

[Ajt06] Miklós Ajtai. Generating random lattices according to the invariant distribution. *Draft of March*, 2006, 2006.

[BCLvV16] Daniel J. Bernstein, Chitchanok Chuengsatiansup, Tanja Lange, and Christine van Vredendaal. Ntru prime. Cryptology ePrint Archive, Report 2016/461, 2016. http://eprint.iacr.org/2016/461.

[BLLN13] Joppe W Bos, Kristin E Lauter, Jake Loftus, and Michael Naehrig. Improved security for a ring-based fully homomorphic encryption scheme. In *IMA Int. Conf.*, pages 45–64. Springer, 2013.

[CJL16] 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. Technical report, Cryptology ePrint Archive, Report 2016/139, 2016.

[CL15] Jung Hee Cheon and Changmin Lee. Cryptanalysis of the multilinear map on the ideal lattices. *IACR Cryptology ePrint Archive*, 2015:461, 2015.

[DDLL13] Léo Ducas, Alain Durmus, Tancrède Lepoint, and Vadim Lyubashevsky. Lattice signatures and bimodal gaussians. In *Advances in Cryptology–CRYPTO 2013*, pages 40–56. Springer, 2013.

[DHS16] Yarkın Doröz, Yin Hu, and Berk Sunar. Homomorphic aes evaluation using the modified ltv scheme. *Designs, Codes and Cryptography*, 80(2):333–358, 2016.

[Gen09] Craig Gentry. Fully homomorphic encryption using ideal lattices. In *STOC*, volume 9, pages 169–178, 2009.

[GGH13] Sanjam Garg, Craig Gentry, and Shai Halevi. Candidate multilinear maps from ideal lattices. In *Eurocrypt*, volume 7881, pages 1–17. Springer, 2013.

[GS02] Craig Gentry and Mike Szydlo. Cryptanalysis of the revised ntru signature scheme. In *Advances in CryptologyEUROCRYPT 2002*, pages 299–320. Springer, 2002.

[HG07] Nick Howgrave-Graham. A hybrid lattice-reduction and meet-in-the-middle attack against ntru. In *Advances in Cryptology-CRYPTO 2007*, pages 150–169. Springer, 2007.

[HHGP$^+$03] Jeffrey Hoffstein, Nick Howgrave-Graham, Jill Pipher, Joseph H Silverman, and William Whyte. Ntrusign: Digital signatures using the ntru lattice. In *Topics in cryptologyCT-RSA 2003*, pages 122–140. Springer, 2003.

[HPS98] Jeffrey Hoffstein, Jill Pipher, and Joseph H Silverman. Ntru: A ring-based public key cryptosystem. In *Algorithmic number theory*, pages 267–288. Springer, 1998.

[HPS11] Guillaume Hanrot, Xavier Pujol, and Damien Stehlé. Terminating bkz. *IACR Cryptology ePrint Archive*, 2011:198, 2011.

[KF15] Paul Kirchner and Pierre-Alain Fouque. An improved bkw algorithm for lwe with applications to cryptography and lattices. In *Advances in Cryptology–CRYPTO 2015*, pages 43–62. Springer, 2015.

[KF17] Paul Kirchner and Pierre-Alain Fouque. Revisiting lattice attacks on overstretched ntru parameters. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 3–26. Springer, 2017.

[LATV12] Adriana López-Alt, Eran Tromer, and Vinod Vaikuntanathan. On-the-fly multiparty computation on the cloud via multikey fully homomorphic encryption. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 1219–1234. ACM, 2012.

[LLL82] Arjen Klaas Lenstra, Hendrik Willem Lenstra, and László Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261(4):515–534, 1982.

[LSS14] Adeline Langlois, Damien Stehlé, and Ron Steinfeld. Gghlite: More efficient multilinear maps from ideal lattices. In *Advances in Cryptology–EUROCRYPT 2014*, pages 239–256. Springer, 2014.

Appendix

A Improve the condition of Theorem 1

Note that the condition of q in **Theorem 1** comes from **Lemma 4**, therefore, we must reduce the size of multiple of $\boldsymbol{g}$ to achieve the better condition. To find smaller $\boldsymbol{a} \cdot \boldsymbol{g}$, applying the algorithm in **Section 3** to several $\boldsymbol{h}_c = [c + \frac{\boldsymbol{f}}{\boldsymbol{g}}]_q = [\frac{\boldsymbol{f}+c\cdot\boldsymbol{g}}{\boldsymbol{g}}]_q$ for small integer c, or adopt other sublattices and other projections. Then we get several $\boldsymbol{a}_1 \cdot \boldsymbol{g}, \cdots \boldsymbol{a}_k \cdot \boldsymbol{g}$ for a sufficiently large k and obtain the ideal lattice generated by $\boldsymbol{g}$ by considering following matrix

$$G = \left(\boldsymbol{a}_1 \cdot \boldsymbol{g}| \cdots |\boldsymbol{a}_k \cdot \boldsymbol{g}\right),$$

because $(\boldsymbol{a}_1) + \cdots (\boldsymbol{a}_k) = R$ might be hold for large k. We let the Hermite normal form HNF(G) of G and consider the sublattice G_l of G which generated by first l column of HNF(G). Then, as [CL15], the lattice reduction algorithm $\mathcal{A}_\epsilon$ with the root Hermite factor ϵ on G_l induces that we get the vector $\boldsymbol{b} = \boldsymbol{a} \cdot \boldsymbol{g}$ for some $\boldsymbol{a} \in R$ such that

$$\|\boldsymbol{b}\| \le \epsilon^l \cdot \|\boldsymbol{g}\|^{n/l}$$

because $\det(G_l) \le \det(G) \le \|\boldsymbol{g}\|^n$. If we choose $l = \sqrt{\frac{n \log M}{\log \epsilon}}$, we finally get $\|\boldsymbol{a} \cdot \boldsymbol{g}\| \le 2^{2\sqrt{n \log M \log \epsilon}}$.

Note that to achieve several $\boldsymbol{a}_i \cdot \boldsymbol{g}$, we must use the lattice reduction algorithm $\mathcal{A}_\delta$ for $\log \delta \le \frac{\log^2 q}{54 n \log \sqrt{2}M}$, that is, the time complexity of this algorithm is still dominant by $\mathcal{A}_\delta$. Choose $\epsilon = \delta$ to get the same asymptotic time complexity to the original algorithm, then one can solve $\mathsf{NTRU}_{R,q,M,q/2}$ by **Lemma 4** if $q^{\frac{2}{3\sqrt{6}}} \le \dfrac{q}{2\|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\|_K \cdot C_F^2}$ holds because $2^{2\sqrt{n \log M \log \epsilon}} \le q^{\frac{2}{3\sqrt{6}}}$.

B Proof of Lemma 6

Lemma 6 Let n be a power of two, and $n = n_t \cdot 2^t$, $\boldsymbol{f}, \boldsymbol{g} \in R = \mathbb{Z}[X]/\langle X^n + 1\rangle$ and $\boldsymbol{g}$ has square free norm $\mathrm{N}_{K/K_t}(\boldsymbol{g})$, $\boldsymbol{h} = [\boldsymbol{f}/\boldsymbol{g}]_q$. If $\boldsymbol{c}_i \in R_t = \mathbb{Z}[X^{2^t}]/\langle X^n + 1\rangle$

for $0 \leq i < 2^t$ satisfies the following inequalities:

$$\|\boldsymbol{c}_i\| < \frac{q}{2^{t+1} \cdot C_F^3 \cdot \|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\|_K \cdot \|\boldsymbol{g}\|^{2^t}} \quad \text{for all} \quad i$$

$$\left\| \left[\sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot \mu_i(\boldsymbol{h}) \right]_q \right\| < \frac{q}{2C_F \cdot \|\boldsymbol{g}\|^{2^t}},$$

then $\boldsymbol{c} = \sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot X^{-i}$ is contained in the ideal $\langle \boldsymbol{g} \rangle$, and the pair $(\boldsymbol{c},\ [\boldsymbol{c} \cdot \boldsymbol{f} \cdot \boldsymbol{g}^{-1}]_q)$ is a solution of $\mathsf{NTRU}_{R,q,M,\frac{q}{2}}$.

Proof. Throughout in this proof, we write $\mathrm{Tr}(\boldsymbol{a})$ and $\mathrm{N}(\boldsymbol{a})$ instead of $\mathrm{Tr}_{K/K_t}(\boldsymbol{a})$ and $\mathrm{N}_{K/Kt}(\boldsymbol{a})$. and define $\tilde{\boldsymbol{h}} = \boldsymbol{f} \cdot \boldsymbol{g}^{-1} \in K$. Trivially, $[\tilde{\boldsymbol{h}}]_q = \boldsymbol{h}$. Let $\boldsymbol{w} := \left[\sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot \mu_i(\tilde{\boldsymbol{h}}) \right]_q$. Since $\mathrm{N}(\boldsymbol{g}) \in R_t$, we get $\mathrm{N}(\boldsymbol{g}) \cdot \mu_i(\tilde{\boldsymbol{h}}) = \mathrm{N}(\boldsymbol{g}) \cdot \mathrm{Tr}(\tilde{\boldsymbol{h}} \cdot X^{-i})/2^t = \mathrm{Tr}(\mathrm{N}(\boldsymbol{g}) \cdot \tilde{\boldsymbol{h}} \cdot X^{-i})/2^t = \mu_i(\mathrm{N}(\boldsymbol{g}) \cdot \tilde{\boldsymbol{h}})$, and following equality holds:

$$[\mathrm{N}(\boldsymbol{g}) \cdot \boldsymbol{w}]_q = \left[\mathrm{N}(\boldsymbol{g}) \cdot \sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot \mu_i(\tilde{\boldsymbol{h}}) \right]_q = \left[\sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot \mu_i(\tilde{\boldsymbol{h}} \cdot \mathrm{N}(\boldsymbol{g})) \right]_q .$$

By two conditions of lemma, we have

1. $\|\mathrm{N}(\boldsymbol{g}) \cdot \boldsymbol{w}\| \leq C_F \cdot \|\mathrm{N}(\boldsymbol{g})\| \cdot \|\boldsymbol{w}\| \leq C_F \cdot \|\boldsymbol{g}\|^{2^t} \cdot \|\boldsymbol{w}\| \leq q/2$
2. $\left\| \sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot \mu_i(\tilde{\boldsymbol{h}} \cdot \mathrm{N}(\boldsymbol{g})) \right\| \leq C_F \cdot \sum_{i=0}^{2^t-1} \|\boldsymbol{c}_i\| \cdot \|\mu_i(\tilde{\boldsymbol{h}} \cdot \mathrm{N}(\boldsymbol{g}))\|$
   $\leq C_F \cdot \sum_{i=0}^{2^t-1} \|\boldsymbol{c}_i\| \cdot \|\tilde{\boldsymbol{h}} \cdot \mathrm{N}(\boldsymbol{g})\| \leq C_F^3 \cdot \sum_{i=0}^{2^t-1} \|\boldsymbol{c}_i\| \cdot \|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\|_K \cdot \|\boldsymbol{g}\|^{2^t} \leq q/2$

Therefore, $\mathrm{N}(\boldsymbol{g}) \cdot \boldsymbol{w} = \sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot \mu_i(\tilde{\boldsymbol{h}} \cdot \mathrm{N}(\boldsymbol{g})) = \sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot \mathrm{Tr}(\tilde{\boldsymbol{h}} \cdot \mathrm{N}(\boldsymbol{g}) \cdot X^{-i})/2^t$ in R_t. It is rewritten as

$$2^t \cdot \mathrm{N}(\boldsymbol{g}) \cdot \boldsymbol{w} - \sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot \big(\mathrm{Tr}(X^{-i} \cdot \boldsymbol{f} \cdot \mathrm{N}(\boldsymbol{g})/\boldsymbol{g}) - X^{-i} \cdot \boldsymbol{f} \cdot \mathrm{N}(\boldsymbol{g})/\boldsymbol{g} \big)$$
$$= (\sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot X^{-i}) \cdot \boldsymbol{f} \cdot \mathrm{N}(\boldsymbol{g})/\boldsymbol{g}.$$

Because the left hand side of the equation is a multiple of $\boldsymbol{g}$ and $\boldsymbol{f} \cdot \mathrm{N}(\boldsymbol{g})/\boldsymbol{g}$ is a relative prime to $\boldsymbol{g}$ by the conditions, we can conclude $\sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot X^{-i} \in \langle \boldsymbol{g} \rangle$.

Moreover, we get

$$\|\sum_{i=0}^{2^t-1} \boldsymbol{c}_i \cdot X^{-i}\| \leq \sum_{i=0}^{2^t-1} \|\boldsymbol{c}_i\| < \frac{q}{2 \cdot C_F^3 \cdot \|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\|_K \cdot \|\boldsymbol{g}\|^{2^t}} < \frac{q}{2C_F^2 \cdot \|\boldsymbol{f}\| \cdot \|\boldsymbol{g}^{-1}\|_K}$$

as a condition. Finally, **Lemma 4** shows that $(\boldsymbol{c},\ [\boldsymbol{c} \cdot \boldsymbol{f} \cdot \boldsymbol{g}^{-1}]_q)$ is a solution of $\mathsf{NTRU}_{R,q,M,\frac{q}{2}}$. □