Fiat-Shamir Signatures based on Module-NTRU

Shi Bai[1], Austin Beard[1], Floyd Johnson[1],
Sulani Kottal Baddhe Vidhanalage[1], and Tran Ngo[1] ⋆

Department of Mathematical Sciences, Florida Atlantic University.

Abstract. Module-NTRU lattices, as a generalization of versatile NTRU lattices, were introduced by Cheon, Kim, Kim and Son (IACR ePrint 2019/1468), and Chuengsatiansup, Prest, Stehlé, Wallet and Xagawa (ASIACCS '20). The Module-NTRU lattices possess the benefit of being more flexible on the underlying ring dimension. They also show how to efficiently construct trapdoors based on Module-NTRU lattices and apply them to trapdoor-based signatures and identity-based encryption. In this paper, we construct Fiat-Shamir signatures based on variant Module-NTRU lattices. Further generalizing Module-NTRU, we introduce the inhomogeneous Module-NTRU problem. Under the assumption that a variation of the search and decisional problems associated with Module-NTRU and inhomogeneous Module-NTRU are hard, we construct two signature schemes. The first scheme is obtained from a lossy identification scheme via the Fiat-Shamir transform that admits tight security in the quantum random oracle model (QROM), following the framework of Kiltz, Lyubashevsky and Schaffner (EUROCRYPT '18). The second scheme is a BLISS-like (Ducas et al., CRYPTO '13) signature scheme based on the search Module-NTRU problem using the bimodal Gaussian for the rejection sampling. At last, we analyze known attacks and propose concrete parameters for the lossy signature scheme. In particular, the signature size is about 4400 bytes, which appears to be the smallest provably secure signature scheme in the QROM achieving 128-bit security.

Keywords: Lattice-based Signature; Module-NTRU Lattice; Fiat-Shamir.

1 Introduction

Lattices have attracted considerable research interest as they can be used to construct efficient cryptographic schemes which are believed to be quantum-resistant. As evidence, many promising candidates submitted to the NIST post-quantum standardization process are based on lattices. Fundamental computational problems in lattice-based cryptography include the Short Integer Solution problem (SIS) [2,35], the Learning With Errors problem (LWE) [40,41,32,11] and the NTRU problem [24,22].

⋆ This work was supported in part by NIST award 60NANB18D216 and by the National Science Foundation under Grant No. 2044855 and 2122229.

Ajtai's seminal work [2] established the worst-to-average connection for the lattice-based primitives based on the SIS problem. It serves as a security foundation for many cryptographic primitives such as hash functions and signatures [2,20,29]. The LWE problem, introduced by Regev [40,41], is extensively used as a security foundation for encryption, signatures and many others [41,20,15,29]. For efficiency, many practical lattice-based cryptosystems are based on assumptions on structured lattices such as the Ring-LWE [32,44], Ring-SIS[33,31,37] and the NTRU problems [23,25]. Introduced by Hoffstein, Pipher and Silverman [23,25], the NTRU assumption is stated informally as follows: given a polynomial h in $R_q := \mathbb{Z}_q[x]/(\phi(x))$, for a cyclotomic polynomial $\phi(x)$ and a positive integer q, where h is the result of dividing one small element by another, find two polynomials $f, g \in R_q$ with small magnitudes such that $h \equiv g/f \pmod{q}$. Following the pioneer work [23,25], the NTRU assumption has been used extensively in various cryptographic constructions such as encryption, signature and many others [22,15,16]. Little is known on the complexity reduction aspects of the NTRU problem (see also [38,36] for progress on this), yet the NTRU assumption with standard parameters remains essentially unbroken after decades of cryptanalysis.

1.1 Previous work

As an important application, SIS/LWE/NTRU problems have been used extensively to obtain post-quantum digital signatures such as [20,29,15,18,10]. There are two main paradigms for constructing practical lattice-based signature schemes in the literature. The first is to use trapdoor sampling algorithms and the hash-and-sign framework, following the work of Gentry, Peikert, and Vaikuntanathan in [20] (GPV). The second framework, proposed by Lyubashevsky [28,29], utilizes the Fiat-Shamir [17] with aborts for transforming identification schemes into signature schemes using variants of SIS/LWE assumptions. We describe related work for both directions.

Computing a short preimage solution for the SIS and ISIS problems has been proven to be as hard as solving certain lattice problems in the worst case [2]. However, with a trapdoor for the matrix $\mathbf{A}$ one can efficiently derive short solutions. In the pioneer work of GPV [20], they show how to efficiently construct a trapdoor for the ISIS problem; more specifically, they give a provable way to sample short solutions without leaking information about the trapdoor. This leads to a natural way for constructing signatures using the hash-then-sign paradigm in the random oracle model (ROM). More efficient trapdoor constructions based on the SIS and LWE problem have been further proposed in [6,34]. These lattice trapdoors require that the trapdoor dimension to be about $m \approx \Theta(n \log q)$ for achieving the optimal trapdoor quality. In work [16], the authors instantiate the GPV framework using the NTRU lattices, which only requires $m = 2n$. It thus leads to a more efficient Identity-Based Encryption (IBE) (and signature scheme). In practice, a power-of-two is usually used for the underlying ring dimension in NTRU, which leads to inflexibility on the parameter selection for desired security level. To overcome such inflexibility, the Module-NTRU (MNTRU) problem was proposed in [13,14] as a generalization of the NTRU problem. The MNTRU takes the equation $\mathbf{F} \cdot \mathbf{h} = \mathbf{g}$, where $\mathbf{h}, \mathbf{g}$ are vectors of polynomials in R_q^{d-1} and $\mathbf{F}$ is an invertible matrix of dimension $d-1$ with elements in R_q. The elements in $\mathbf{F}, \mathbf{g}$ are small for the MNTRU problem to be well-defined. The work [13,14] constructed trapdoors and proposed instantiations of the hash-then-sign paradigm using the MNTRU assumption. Concrete instantiations of the hash-then-sign signatures include the NIST PQC submissions Falcon [39], pqNTRUSign [45], etc.

The signatures discussed above use the trapdoor functions with the hash-and-sign paradigm. A second paradigm to construct lattice-based signatures is to use the Fiat-Shamir transform [17]. In [28,29], Lyubashevsky utilizes Fiat-Shamir for transforming identification schemes into provably secure signature schemes using variants of SIS/LWE assumptions. In particular, the rejection sampling in Fiat-Shamir is proposed to ensure the distribution of the signatures is independent from the private key and hence preventing the leakage of private keys. An improvement, the so-called BLISS scheme [15], is obtained by using the bimodal Gaussian distribution in the rejection sampling. This leads to a much smaller rejection area for signatures. For practical instantiation, BLISS [15] also devised an efficient signature scheme using the NTRU assumption. Follow-up work such as [21,15,7,5] uses a compression technique to further reduce the signatures size: the common idea is to throw away some bits of the vector to be hashed. The security proofs in these works remain non-tight due to the use of the Forking Lemma [8] with the reprogramming of random oracles. Furthermore, their security is usually studied in the random oracle model.

To construct signature schemes with tight security, Abdalla, Fouque, Lyubashevsky and Tibouchi [1] proposed the lossy identification scheme, and proved that the signatures obtained from Fiat-Shamir admit a tight security in the ROM model. A similar approach has been used in the TESLA signature scheme [5,4]. The general idea is to start with a lossy identification scheme which adopts two security properties, e.g. key indistinguishability and lossiness: it admits a lossy key generation algorithm that produces a lossy public key which is computationally indistinguishable to the genuine public keys, yet it is statistically impossible to win the impersonation game when the public key is lossy. The signature derived from such an identification scheme [1] was known to be secure only in the random oracle model, which does not automatically imply security in the quantum random oracle model (QROM). Kiltz, Lyubashevsky and Schaffner [26] presented a generic Fiat-Shamir framework from lossy identification schemes [1] to obtain tight secure signatures in the QROM. By adaptively re-programming of the random oracle, the same tight security result in the QROM has been obtained for the TESLA signature scheme [5,4]. A concrete instantiation of [26] is to adapt and to modify the Dilithium signature scheme [30], which has tight secure reductions from Module-SIS (MSIS) and Module-LWE (MLWE). A concrete instantiation of the techniques in [5,4] is given in the qTESLA signature [10], whose existential unforgeability under chosen message attack (EUF-CMA) security is reduced from the underlying decisional Ring-LWE problem.

To our knowledge, the minimum signature size that achieves near 128-bit security in the QROM model is from [26] with a pair of parameter sets given. The first set has a signature size of 5690 bytes and public key size 7712 bytes whose public key prevents a BKZ reduction of block size up to 480. The second set admits a larger key security (BKZ block size of 600) has signature size 7098 bytes and public key size 9632 bytes.

1.2 Contributions

In this work we present two Fiat-Shamir signature schemes based on some variant Module-NTRU problems. The first scheme follows the framework of [26], starting from an identification scheme and applying the Fiat-Shamir transform. The second scheme is analogous to the BLISS [15] scheme, but built on the variant Module-NTRU problem, with a fixed q being part of the public key. Thus, they may be viewed as variants of the signatures from [26] and BLISS [15], instantiated with the (inhomogeneous) Module-NTRU assumptions.

We first generalize the Module-NTRU problem proposed in [13,14] to the inhomogeneous MNTRU (iMNTRU) problem and formalize the hardness assumptions used. Briefly, the iMNTRU consists of the equation $\mathbf{F} \cdot \mathbf{h} + \mathbf{g} = \mathbf{t}$, where $\mathbf{t}$ comes from a certain distribution. In our signature, essentially the $\mathbf{F}$ and $\mathbf{g}$ serve as small secrets, while the $\mathbf{h}$ and $\mathbf{t}$ are public keys. The first signature scheme follows the lossy key identification paradigms of [26] using a uniform distribution for nonce generation. We prove the identification scheme achieves completeness of normal keys, simulatability of transcripts, lossy keys, sufficient entropy and computational unique response properties, thus possessing a tight security in the quantum random oracle model to the inhomogeneous Module-NTRU problem. Our second construction is a signature scheme based on the variant MNTRU assumption with a fixed q being part of the public key, and with the bimodal Gaussian distribution. The construction follows a similar framework as the BLISS signature [15], but uses the variant MNTRU assumption, which admits extra flexibility in the choice of parameters for the underlying ring dimension. With these proposed schemes, we analyze known attacks and their efficacy.

We discuss several related works. In [19], Genise at al. described inhomogeneous variants of NTRU problem named MiNTRU. In matrix form, the problem is defined as $\mathbf{A} := \mathbf{S}^{-1}(\mathbf{G} - \mathbf{E}) \pmod{q}$ where $\mathbf{G}$ is a gadget matrix of the form $\mathbf{G} = (\mathbf{0} \mid \mathbf{I} \mid 2\mathbf{I} \mid \cdots \mid 2^{\log q - 1}\mathbf{I})$. The secret matrices $\mathbf{S}$ and $\mathbf{E}$ are sampled from distributions of small magnitudes and the search MiNTRU problem asks an adversary to recover $\mathbf{S}$ and $\mathbf{E}$ from $\mathbf{A}$. In this paper, we introduce a somewhat different assumption by sampling uniformly a vector of polynomials $\mathbf{t} \in R_q^{d-1}$, an invertible matrix of small polynomials $\mathbf{F} \in R_q^{(d-1)\times(d-1)}$ and a vector of small polynomials $\mathbf{g} \in R_q^{d-1}$ so that $\mathbf{h} = \mathbf{F}^{-1}(\mathbf{t} - \mathbf{g})$. In our second BLISS-like signature scheme, we also consider the case where $\mathbf{t}$ is pre-fixed. A work from Chen, Genise and Mukherjee [12] introduced the *approximated* ISIS trapdoor and used it to construct signatures using the hash-and-sign framework, which resulted in reduced sizes on the trapdoor and signature from [34]. For certain distributions, the approximate ISIS problem is shown to be as hard as the standard ISIS problem. The approximate ISIS problem of a given matrix $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$ and a vector $\mathbf{y} \in \mathbb{Z}_q^n$ asks to find a short vector $\mathbf{x}$ from $\mathbb{Z}_q^m$ so that $\mathbf{A}\,\mathbf{x} = \mathbf{y} + \mathbf{z}$ where $\mathbf{z}$ is a small shift. Note the public matrix $\mathbf{A}$ is drawn uniformly, while in our iMNTRU the public vector $\mathbf{h}$ is computed as $\mathbf{h} = \mathbf{F}^{-1}(\mathbf{t} - \mathbf{g})$. Thus, when $\mathbf{F}$ and $\mathbf{g}$ consist of sufficiently small polynomials, the distribution $(\mathbf{h}, \mathbf{t})$ cannot be uniform, yet depending on the distribution of $\mathbf{t}$, the marginal distribution of $\mathbf{h}$ might be uniform.

Existing signature schemes built on the Fiat-Shamir paradigms such as Dilithium [30] and qTESLA [10] are quite efficient and practical. Our scheme further optimizes the scheme parameters such as the signature size. In particular, we achieve a 128-bit security with a signature size of 4400 bytes and a public key size of 10272 bytes for BKZ block size 490. This appears to be smallest provably secure signature scheme in the QROM achieving 128-bit security. We also have a signature size of 9264 bytes and a public key size of 18464 bytes for BKZ block size 669. In addition to parameter optimization, we think it is also beneficial to investigate a more diverse selection of the underlying hardness assumption. One notes that the schemes [30] and qTESLA [10] are both built on the Module-LWE assumptions.

Finally, compared to the BLISS signature [15], the use of the Module-NTRU enjoys the extra flexibility in the choice of parameters for the underlying ring dimension, since many applications require the NTRU lattice to be defined on the power-of-two cyclotomic rings. Thus, sometimes when a higher security level is needed, the dimension of the NTRU lattice needs to be doubled. Recent progress on the complexity aspects of the NTRU problem [38] may shed light on the hardness of the inhomogeneous Module-NTRU problem used in this work.

2 Preliminaries

We present the notation and definitions used to construct our signatures. Let q be an integer, which is usually a prime in this paper. Let $\mathbb{Z}_q$ be the set of all integers modulo q in the range $(-\frac{q}{2}, \frac{q}{2}]$ when q is even and $[-\lfloor\frac{q}{2}\rfloor, \lfloor\frac{q}{2}\rfloor]$ when q is odd. We will refer to it as the *balanced representation mod q*. We denote R and R_q as the rings $\mathbb{Z}[x]/(x^n + 1)$ and $\mathbb{Z}_q[x]/(x^n + 1)$, respectively. The integer n is usually a power of 2, where $q \equiv 1 \pmod{2n}$. In this case, the polynomial $X^n + 1$ splits completely in $\mathbb{Z}_q$. Throughout, regular font letters such as v denote ring elements in R, R_q and $\mathbb{Z}, \mathbb{Z}_q$. We use bold lower-case letters such as $\mathbf{v}$ to represent vectors of elements from their respective fields. For a vector $\mathbf{v}$, we denote by $\mathbf{v}^{\mathbf{t}}$ its transpose, we also denote $\mathbf{0}$ to be the zero vector. Bold upper case letters denote matrices. A matrix $\mathbf{B} = (\mathbf{b}_1, \cdots, \mathbf{b}_n)$ is also presented in a column-wise way. Abusing notation, we sometimes also use lower-case letters to identify the coefficients of ring elements in R and R_q.

For a polynomial $f = \sum_{i=0}^{n-1} a_i x^i \in R_q$, we identify its *coefficient embedding* as its vector of coefficients $f := (a_0, \ldots, a_{n-1})^T$. For a vector of polynomials $\mathbf{f} = (f_1, \ldots, f_n) \in R_q^n$, we may use $v_{\mathbf{f}}$ as a coefficient vector $(f_1, \ldots, f_n)^T$. A polynomial f in R_q can be associated with an acyclic matrix M_f. Multiplying $f(x)$ by $g(x) = \sum_{i=0}^{n-1} g_i x^i \in R_q$ identifies with the product of $M_f \cdot \mathbf{g}$. For a vector $\mathbf{x}$, we use $\|\mathbf{x}\|$ to denote its ℓ_2-norm and $\|\mathbf{x}\|_\infty = \max_i(|\mathbf{x}_i|)$ to denote its ℓ_∞-norm. The ℓ_2-norm and ℓ_∞-norm of polynomial f are defined as the corresponding norms on the corresponding coefficient vector. Given a vector $\mathbf{f}$ consisting of polynomials f_i, the norm notation extends naturally, i.e., $\|\mathbf{f}\|_\infty = \max_i(\|f_i\|_\infty)$. The inner product of two vectors $\mathbf{x}$ and $\mathbf{y}$ is denoted by $\langle \mathbf{x}, \mathbf{y} \rangle$. For convenience, we define some notations for rounding.

For an integer $c \in \mathbb{Z}$, we denote $[c]_r$ to be the unique integer in the range $(-2^{r-1}, 2^{r-1}]$ such that $[c]_r \equiv c \pmod{2^r}$. We denote $c = \lfloor c \rceil_r \cdot 2^r + [c]_r$, where $\lfloor c \rceil_r$ extracts the higher bits of c. In this paper, the inputs c will be in balanced representation mod q. For a polynomial $f = \sum_{i=0}^{n-1} a_i x^i$ we extend $[.]_r$ and $\lfloor . \rceil_r$ to f on its coefficients coordinate-wise. We define $\mathcal{B}_{n,\kappa}$ to be the set of ternary (or binary) vectors of length n with Hamming weight κ. When the length n is clear in the context, we may write $\mathcal{B}_\kappa$ for short.

We will use the rejection sampling lemma from [29] to ensure the output signature does not leak information about the secret key. We review the definition of various distributions and rejection sampling lemma, and the background on lattices (see full version of this work). We also review the background on identification, digital signatures and the Fiat-Shamir transform (see full version of this work).

2.1 (Inhomogeneous) Module-NTRU

As a generalization of NTRU, the Module-NTRU (MNTRU) problem was introduced in [13,14], which enables the dimension and parameter flexibility. It was used to construct trapdoors for lattice signatures and identity-based encryption (IBE). Intuitively, given a vector $\mathbf{h}$ such that the inner product of $(1, \mathbf{h})$ and some "small" secret vector $\mathbf{f}$ is zero, the Module-NTRU problem asks to recover the secret $\mathbf{f}$ or close. In this paper, we will use a natural variant of the Module-NTRU, which we denote as the *inhomogeneous* Module-NTRU (iMNTRU) problem. We formalize the problem as follows.

Definition 1 **($\mathsf{iMNTRU}_{q,n,d,B}$ instance).** *Let $n, d \geq 2$ be integers, and q be a prime. Let B be a positive real number. Denote $R_q = \mathbb{Z}_q[x]/(x^n + 1)$. An $\mathsf{iMNTRU}_{q,n,d,B}$ instance consists of a vector $\mathbf{h} \in R_q^{d-1}$ and $\mathbf{t} \in R_q^{d-1}$ such that there exists an invertible matrix $\mathbf{F} \in R_q^{(d-1)\times(d-1)}$ and a vector $\mathbf{g} \in R_q^{d-1}$ with $\mathbf{F} \cdot \mathbf{h} + \mathbf{g} = \mathbf{t} \pmod q$ and $\|\mathbf{F}\|, \|\mathbf{g}\| \leq B$. The $(\mathbf{F}, \mathbf{g})$ is called a trapdoor of the $\mathsf{MNTRU}_{q,n,d,B}$ instance $\mathbf{h}$. An $\mathsf{MNTRU}_{q,n,d,B}$ instance corresponds to an $\mathsf{iMNTRU}_{q,n,d,B}$ instance for the case when $\mathbf{t} = \mathbf{0}$.*

Definition 2 **($\mathsf{iMNTRU}_{q,n,d,D_1,D_2,T}$ distribution).** *Let n, d be positive integers, and q be a prime. Let D_1, D_2, T be distributions defined over $R_q^{(d-1)\times(d-1)}$, R_q^{d-1} and R_q^{d-1} respectively. An $\mathsf{iMNTRU}_{q,n,d,D_1,D_2,T}$ sampler is a polynomial-time algorithm that samples matrix $\mathbf{F}$ from D_1, vector $\mathbf{g}$ from D_2, vector $\mathbf{t}$ from T and* *then computes* $\mathbf{h}$ *in* $\mathbf{F}\cdot\mathbf{h}+\mathbf{g}=\mathbf{t} \pmod q$*. The sampler outputs a tuple* $(\mathbf{h},\mathbf{F},\mathbf{g},\mathbf{t})$*. An* $\mathsf{iMNTRU}_{q,n,d,D_1,D_2,T}$ *distribution is the induced marginal distribution of* $(\mathbf{h},\mathbf{t})$ *from an* $\mathsf{iMNTRU}_{q,n,d,D_1,D_2,T}$ *sampler. For the distribution to be meaningful, we usually assume* D_1, D_2 *are* B*-bounded distributions and* D_1 *turns out to be an distribution defined on invertible elements* $\mathbf{F}$*. An* $\mathsf{MNTRU}_{q,n,d,D_1,D_2}$ *distribution corresponds to the case of an* $\mathsf{iMNTRU}_{q,n,d,D_1,D_2,T}$ *distribution when the support of* T *is always* 0*.*

In the schemes presented in this work, we will make several different choices for the distribution T, depending on the design and functionality. The decisional variant and search variant of the MNTRU are defined as follows:

Definition 3 (Decisional $\mathsf{iMNTRU}_{q,n,d,D_1,D_2,T,B}$). *Let* n, d *be positive integers, and* q *be a prime. Let* D_1, D_2 *be* B*-bounded distributions defined over* $R_q^{(d-1)\times(d-1)}$ *and* R_q^{d-1} *respectively, and* T *be a distribution over* R_q^{d-1}*. Let* $\mathcal{N}$ *be an* $\mathsf{iMNTRU}_{q,n,d,D_1,D_2,T}$ *distribution. The decisional* $\mathsf{iMNTRU}_{q,n,d,D_1,D_2,T,B}$ *problem asks to distinguish between samples from* $\mathcal{N}$ *and from* $U(R_q^{d-1})\times T$*. The decisional* $\mathsf{MNTRU}_{q,n,d,D_1,D_2,B}$ *is defined similarly when the support of* T *is always* 0*.*

Definition 4 (Search $\mathsf{iMNTRU}_{q,n,d,D_1,D_2,T,B}$). *Let* n, d *be positive integers, and* q *be a prime. Let* D_1, D_2 *be* B*-bounded distributions defined over* $R_q^{(d-1)\times(d-1)}$ *and* R_q^{d-1} *respectively, and* T *be a distribution over* R_q^{d-1}*. Let* $\mathcal{N}$ *denote the* $\mathsf{iMNTRU}_{q,n,d,D_1,D_2,T,B}$ *distribution. Given samples* $(\mathbf{h},\mathbf{t})$ *from* $\mathcal{N}$*, the search* $\mathsf{iMNTRU}_{q,n,d,D_1,D_2,T,B}$ *problem is to recover an invertible* $\mathbf{F}$ *and* $\mathbf{g}$ *such that* $\mathbf{F}\cdot\mathbf{h}+\mathbf{g}=\mathbf{t} \pmod q$ *and* $\|\mathbf{F}\|, \|\mathbf{g}\| \leq B$*. The search* $\mathsf{MNTRU}_{q,n,d,D_1,D_2,B}$ *is defined similarly when the support of* T *is always* 0*. Given an* $\mathsf{iMNTRU}_{q,n,d,B}$ *instance* $(\mathbf{h},\mathbf{t})$*, the worst-case search* $\mathsf{iMNTRU}_{q,n,d,B}$ *problem is to recover an invertible* $\mathbf{F}$ *and* $\mathbf{g}$ *such that* $\mathbf{F}\cdot\mathbf{h}+\mathbf{g}=\mathbf{t} \pmod q$ *and* $\|\mathbf{F}\|, \|\mathbf{g}\| \leq B$*. The worst-case search* $\mathsf{MNTRU}_{q,n,d,B}$ *problem is defined when* $\mathbf{t}$ *is* 0*. Clearly, the worst-case search* $\mathsf{MNTRU}_{q,n,d,B}$ *problem reduces to worst-case search* $\mathsf{iMNTRU}_{q,n,d,B}$ *problem.*

We are not aware of any reduction between MNTRU and the average cases of inhomogeneous MNTRU assumptions where the $\mathbf{t}$ is sampled from a distribution. However, one can reduce from MNTRU to inhomogeneous MNTRU by assuming a worst-case oracle on the inhomogeneous MNTRU problem. We will make the assumption that the average-case inhomogeneous MNTRU assumption is as hard as the MNTRU assumption. Our signature scheme relies on an additional assumption that solving a single row of the iMNTRU assumption is as hard as the iMNTRU assumption. Namely, our signature schemes only use a single row $\mathbf{f}$ of $\mathbf{F}$ and hence the vectors $\mathbf{g}, \mathbf{t}$ are just two polynomials, thus the equation becomes $\langle \mathbf{h}, \mathbf{f}\rangle + g = t \pmod q$. The variant search and decisional problems are defined correspondingly and we require that $\mathbf{f}$ is non-zero.

Our first signature scheme reduces from this variant search and decisional inhomogeneous Module-NTRU assumptions, which we assumed hard to invert and indistinguishable from uniform respectively. Our second signature scheme is based on the variant search Module-NTRU assumption, which is assumed hard to invert as in [13,14].

Remark 1. In the key generation presented in this work, one actually just starts with a single vector $\mathbf{f}$ and pick up an element $\mathbf{h}$ in the left kernel of $t - g$ w.r.t. $\mathbf{f}$. One can pick up $\mathbf{h}$ by choosing h_i for $i \leq d-2$ first and then computing h_{d-1} in the end. We note here that the distributions of the public keys for our assumption and iMNTRU are not the same. We will make the assumption that this variant assumption is as hard as the iMNTRU assumption. This variant assumption turns out to be analogous to "low-density" inhomogeneous Ring-SIS problem [29]. We leave for future work to study its average-case hardness.

3 Signature based on iMNTRU in the QROM

In this section, we present a lossy identification scheme based on the variant of inhomogeneous Module-NTRU assumption. Our construction follows the design and paradigm proposed in [1,4,26] via the Fiat-Shamir transformation and thus leads to a tightly-secure signature in the quantum random-oracle model. In this work, the random oracle H takes inputs from $R_q \times \mathcal{M}$, where $\mathcal{M}$ denotes the message space, and outputs a polynomial in R_q. We restrict the output polynomials to be ternary (or binary) and have κ non-zero coefficients, e.g. those can be identified as vectors from $\mathcal{B}_{n,\kappa}$. We refer to [15] for efficient instantiation of random oracles.

3.1 A lossy identification scheme

As in [1,26], we start by constructing a lossy identification scheme ID, given in Figure 1. The key generation algorithm starts by choosing parameters $d \in \mathbb{N}$ as the rank, n as the ring dimension and a prime q as the modulus. Similar to the key generation of [13,14], one can sample $(\mathbf{h}', \mathbf{F}, \mathbf{g}, \mathbf{t})$ from an $\mathsf{iMNTRU}_{q,n,d,D_1,D_2,U(R_q)}$ distribution, where D_1 and D_2 are two distributions for sampling the secret keys. Here we sample each f in $\mathbf{F}$ from U_β^n and each g in $\mathbf{g}$ from U_β^n independently. Note that it is possible to sample them from other "small" distributions such as discrete Gaussian, but we use uniform distribution here. After we sample $\mathbf{g}, \mathbf{t}$ and an invertible $\mathbf{F}$, we compute $\mathbf{h}' = \{h_i\}_{i=1}^{d-1}$ in $\mathbf{F} \cdot \mathbf{h}' + \mathbf{g} = \mathbf{t} \pmod{q}$. Note that for cryptographically sized parameters the probability that a randomly selected matrix of polynomials $\mathbf{F}$ is invertible is close to one.

As previously mentioned, one can only use a single row $(f_1, \ldots, f_{d-1})$ from $\mathbf{F}$ and let g, t be corresponding polynomials in $\mathbf{g}, \mathbf{t}$, respectively. Abusing notation, we denote $f_d := g$ and $\mathbf{f} = (f_1, \ldots, f_{d-1}, f_d)$, which is the secret key for our identification scheme. We also denote $\mathbf{h} = (h_1, \ldots, h_{d-1}, 1)$ and set $(\mathbf{h}, t)$ as the public key. With this rewrite, we see that $\langle \mathbf{h}, \mathbf{f} \rangle = t$. We use balanced representation mod q in the following algorithm.

In the first step of the identification, the prover samples a vector of polynomials $\mathbf{y} := (y_1, \ldots, y_d)$, where each y_i is from the distribution U_γ^n, and computes the

Algorithm $\mathsf{IGen}(q, n, d, \beta)$

```
Sample $\mathbf{f} = \{f_i\}_{i=1}^{d}$ and $t$, where $f_i \hookleftarrow U_\beta^n$ and $t \hookleftarrow U_{R_q}$
Compute $\mathbf{h} = (h_1, \dots, h_{d-1}, 1)$ such that $\sum_{i=1}^{d} h_i f_i \equiv t \pmod q$
return $\mathsf{pk} := (\mathbf{h}, t)$ and $\mathsf{sk} := \mathbf{f}$
```

Algorithm $\mathsf{P}_1(\mathsf{sk})$:

```
Sample $\mathbf{y} = \{y_i\}_{i=1}^{d-1}$ where $y_i \hookleftarrow U_\gamma^n$
Compute $u = \left\lfloor \sum_{i=1}^{d-1} h_i y_i \pmod q \right\rceil_r$
return $u$
```

Algorithm $\mathsf{P}_2(\mathsf{sk}, u, c)$:

```
Compute $\mathbf{z} := (z_1, \dots, z_{d-1})$ where $z_i = y_i + c \cdot f_i$
Compute $w = \sum_{i=1}^{d-1} h_i y_i - c \cdot f_d \pmod q$
if any $\|z_i\|_\infty > \gamma - \beta \cdot \kappa$
  or $\|[w]_r\|_\infty \geq 2^{r-1} - \beta \cdot \kappa$
  or $\|w\|_\infty \geq \lfloor q/2 \rfloor - \beta \cdot \kappa$ then
    return $\perp$
return $\mathbf{z}$
```

Algorithm $\mathsf{V}(\mathsf{pk}, u, c, \mathbf{z})$:

```
if $\forall 1 \leq i \leq d-1, \|z_i\|_\infty \leq \gamma - \beta \cdot \kappa$ and $\left\lfloor \sum_{i=1}^{d-1} h_i z_i - t \cdot c \pmod q \right\rceil_r = u$ then
    return Accept
return Reject
```

Fig. 1. A lossy identification scheme based on variant of iMNTRU

commitment $u := \left\lfloor \sum_{i=1}^{d-1} h_i y_i \pmod q \right\rceil_r$. The prover then sends u to the verifier. The verifier generates a random challenge c from the distribution $\mathcal{B}_\kappa$ (here we define it to be the set of ternary vectors of length n with weight κ) and sends c to the prover. The number of nonzero coefficients in c is κ, thus the infinity norm of $f_i \cdot c$ is bounded by $\beta \cdot \kappa$. The prover computes $z_i := y_i + c \cdot f_i$ and returns $\mathbf{z}$ if, for all $1 \leq i \leq d-1$, $\|z_i\|_\infty \leq \gamma - \beta \cdot \kappa$, and $|[\sum_{i=1}^{d-1} h_i y_i - c \cdot f_d \pmod q)]_r| < 2^{r-1} - \beta \cdot \kappa$ together with $\|w\|_\infty < \lfloor q/2 \rfloor - \beta \cdot \kappa$. Otherwise, it returns $\perp$. Verifier accepts $(\mathbf{z}, u)$ if, for all i, we have $\|z_i\|_\infty \leq \gamma - \beta \cdot \kappa$ and $\left\lfloor \sum_{i=1}^{d-1} h_i z_i - t \cdot c \pmod q \right\rceil_r$ equals u. Otherwise, it rejects. To optimize slightly, it is possible to record $\sum_{i=1}^{d-1} h_i y_i$ as a state for the prover in Algorithm P_1 and re-use in Algorithm P_2.

In this section, we present the lossy identification scheme in Figure 1. We show the scheme admits properties including na-HVZK, correctness, lossy, min-entropy and computational unique response (CUR). The proof follows a similar framework as in [26]. For Lemmas 1 to 5, we state them and sketch the proofs in the full version of this work.

We first show that the ID scheme is perfectly na-HVZK. Following the definition of na-HVZK, we set two algorithms Sim(.) and Trans(.), shown in Figure 2. We will show that the distribution of outputs of Sim(.) and Trans(.) is identical. For convenience, we denote $B := \beta \cdot \kappa$.

Lemma 1. *The identification scheme of Figure 1 is perfect* na-HVZK*.*

We now prove that the identification is correct, up to some rejection rate. We stress that such a bound is not rigorous, as we assumed a specific distribution on the rounded numbers, yet it is sufficient to use in practice. One can get a more accurate rejection rate from a simulation.

Lemma 2. *Under the variant decisional iMNTRU assumption, the identification scheme has correctness error*

$$\delta \approx 1 - \exp\left(-\beta\kappa n\left(\frac{d-1}{\gamma} + \frac{1}{2^{r-1}} + \frac{1}{q}\right)\right).$$

We now show that the identification scheme is lossy. We first define a lossy key generation algorithm $\mathsf{LossyIGen}(q, n, d, \beta)$, shown in Figure 3, which samples h_i's and t from uniform. First, the public keys generated by LossyIGen and IGen are indistinguishable due to the variant decisional iMNTRU assumption. It remains to show the scheme admits $\varepsilon_{\mathsf{ls}}$-lossy soundness; that is, for any quantum adversary, the probability of impersonating the prover is bounded by $\varepsilon_{\mathsf{ls}}$.

Lemma 3. *The identification scheme admits ϵ_{ls}-lossy soundness for*

$$\epsilon_{\mathsf{ls}} \leq \frac{1}{|\mathcal{B}_\kappa|} + 2 \cdot |\mathcal{B}_\kappa|^2 \cdot \frac{(4(\gamma - B) + 1)^{n(d-1)} \cdot (2^{r+1} + 1)^n}{q^n}.$$

```
Algorithm $\mathsf{Trans}(\mathsf{sk})$
    Sample $\mathbf{y} = \{y_i\}_{i=1}^{d-1}$ where $y_i \hookleftarrow U_\gamma^n$
    Compute $u = \left\lfloor \sum_{i=1}^{d-1} h_i y_i \pmod q \right\rceil_r$
    Sample $c \hookleftarrow \mathcal{B}_\kappa$
    Compute $\mathbf{z} = \{z_i\}_{i=1}^{d-1}$ where $z_i = y_i + c \cdot f_i$
    Compute $w = \sum_{i=1}^{d-1} h_i y_i - c \cdot f_d \pmod q$
    if any $\|z_i\|_\infty > \gamma - B$ return $\perp$
    if $\left\| [w]_r \right\|_\infty \geq 2^{r-1} - B$
      or $\|w\|_\infty \geq \lfloor q/2 \rfloor - B$ return $\perp$
    return $(\mathbf{z}, c)$

Algorithm $\mathsf{Sim}(\mathsf{pk})$
    With probability $1 - \left( \frac{|U_{\gamma-B}|}{|U_\gamma|} \right)^{n(d-1)}$
        return $\perp$
    Sample $\mathbf{z} = \{z_i\}_{i=1}^{d-1}$ where $z_i \hookleftarrow U_{\gamma-B}^n$
    Sample $c \hookleftarrow \mathcal{B}_\kappa$
    Compute $w' = \sum_{i=1}^{d-1} h_i z_i - t \cdot c \pmod q$
    if $\left\| [w']_r \right\|_\infty \geq 2^{r-1} - B$
      or $\|w'\|_\infty \geq \lfloor q/2 \rfloor - B$ return $\perp$
    return $(\mathbf{z}, c)$
```

Fig. 2. Transcript algorithm and simulation algorithm

```
Algorithm $\mathsf{LossyIGen}(q, n, d, \beta)$
    Sample $\mathbf{h} = (h_1, \ldots, h_{d-1}, 1)$ and $t$, where $h_i \hookleftarrow U_{R_q}$ and $t \hookleftarrow U_{R_q}$
    return $\mathsf{pk} := (\mathbf{h}, t)$
```

Fig. 3. Lossy key generation algorithm $\mathsf{LossyIGen}$

This bound essentially says q should be larger than γ^d asymptotically. This condition is natural, since otherwise, it is intuitive to see there exist many solutions $\mathbf{z}, c$ for $u = \left\lfloor \sum_{i=1}^{d-1} h_i z_i - t \cdot c \right\rceil_r$.

We now prove that the u sent by the prover in Algorithm P_1 is very likely to be distinct across every run of the protocol. We first remark that the public key $\mathbf{h}' \hookleftarrow \mathsf{IGen}$ (i.e. recall that $\mathbf{h} = (\mathbf{h}', 1)$) has a marginal distribution which is uniform in R_q^{d-1}. This is because $\mathbf{h}'$ is computed in equation $\mathbf{F} \cdot \mathbf{h}' + \mathbf{g} = \mathbf{t} \pmod q$ where $\mathbf{t}$ is uniform and $\mathbf{F}$ is invertible. Note that the joint distribution $(\mathbf{h}', \mathbf{t})$ is not uniform for our choice of parameters, but in Algorithm P_1, only $\mathbf{h}'$ is used to produce the commitment.

Lemma 4. *The identification scheme has $\alpha := n \cdot \log E$ bits of min-entropy, where*

$$E = \min \left\{ (2\gamma + 1)^{d-1},\ \frac{q}{(4\gamma + 1)^{(d-1)}(2^{r+1} + 1)} \right\}.$$

In the end, we sketch that our scheme satisfies the computational unique response (CUR) property for the strong unforgeability of the signature scheme after the Fiat-Shamir transform.

Lemma 5. *For any adversary on the identification scheme, the success probability of producing two valid transcripts $(u, c, \mathbf{z})$ and $(u, c, \mathbf{z}')$, such that $\mathbf{z} \neq \mathbf{z}'$, is bounded by $(4(\gamma - B) + 1)^{n(d-1)} \cdot (2^{r+1} + 1)^n \cdot q^{-n}$.*

In the end, we give the signature scheme constructed from the lossy identification scheme (see full version of this work). Theorem 3.1 of [26] concludes that the signature scheme admits a tight security in the QROM. The concrete parameters for the signature scheme will be given in Section 5.1.

4 A BLISS-like signature based on MNTRU

In this section, we propose a signature scheme based on the variant MNTRU assumption with a fixed t and the bimodal Gaussian distribution. The construction follows a similar framework as the BLISS signature [15], but uses the variant MNTRU assumption, which admits the extra flexibility in the choice of parameters for the underlying ring dimension.

4.1 Signature scheme

We give the signature scheme in Figure 4 and describe the key generation, signing and verification procedure here. In Algorithm Gen, used for key generation, one chooses the following parameters: rank $d \in \mathbb{N}$, a prime modulus q, an integer n as the ring dimension, and a positive odd integer $\beta < q$. We sample $(\mathbf{h}', \mathbf{F}, \mathbf{g})$ from the $\mathsf{MNTRU}_{q,n,d,D_1,D_2}$ distribution, where D_1 is U_β^n and D_2 is $U_{\lfloor \beta/2 \rfloor}^n$ are distributions of secret keys $\mathbf{F}$ and $\mathbf{g}$, respectively. It is sufficient to take a single row $\{f_i\}_{i=1}^{d-1}$ from $\mathbf{F}$ and we denote $\mathbf{s} = (f_1, \ldots, f_{d-1}, f_d)$ where

$f_d := 2g+1$. Note the coefficients of f_d also lie uniformly in $[-\beta, \beta]$. We denote $\mathbf{h} = (h_1, \ldots, h_{d-1}, -1)$ and hence $\langle \mathbf{h}, \mathbf{s} \rangle = 0 \pmod q$. In the scheme, we use the vector $\mathbf{a} = (2h_1, \cdots, 2h_{d-1}, q-2) \in R_{2q}^d$ as the public key and vector $\mathbf{s} \in R_{2q}^d$ as the private key. It can be checked that we have $\langle \mathbf{a}, \mathbf{s} \rangle \equiv q \pmod{2q}$, since

$$\langle \mathbf{a}, \mathbf{s} \rangle \equiv \sum_{i=1}^{d-1} 2h_i f_i - 2f_d \equiv 0 \pmod q,$$
$$\langle \mathbf{a}, \mathbf{s} \rangle \equiv q \cdot (2g+1) \equiv 1 \pmod 2.$$

To sign a message μ, the signer chooses a vector $\mathbf{y} := (y_1, \ldots, y_d)$, where each y_i is sampled from the discrete Gaussian $D_{\mathbb{Z},\sigma}^n$. The signer then computes $c := H(\langle \mathbf{a}, \mathbf{y} \rangle \pmod{2q}, \mu)$ and $\mathbf{z} := \mathbf{y} + (-1)^b c \cdot \mathbf{s}$ for a uniform random bit $b \in \{0,1\}$. With rejection sampling, the signature $(c, \mathbf{z})$ is outputted with probability $1/M \exp(-\|c \cdot \mathbf{s}\|^2/(2\sigma^2)) \cosh(\langle \mathbf{z}, c \cdot \mathbf{s} \rangle / \sigma^2)$, where the constant M is the repetition rate for each signing. Upon receiving the signature $(c, \mathbf{z})$, the verification will succeed if $\|\mathbf{z}\|_\infty < q/4$, $\|\mathbf{z}\| \leq \eta\sigma\sqrt{nd}$, and $H(\langle \mathbf{a}, \mathbf{z} \rangle + q \cdot c \pmod{2q}, \mu) = c$. For convenience, we did not use compression in the presented scheme, but mention it should be similar to [15] to compress the signature.

Rejection Sampling. The rejection sampling follows the same as [15]. Consider $\mathbf{z} = (-1)^b \cdot \mathbf{s} \cdot c + \mathbf{y}$. Abusing notation, we denote $\mathbf{s} \cdot c$ as the concatenated coefficient vector as well as a vector of polynomials. The distribution of $\mathbf{z}$ is the bimodal discrete Gaussian distribution $\frac{1}{2} D_{\mathbb{Z}^{nd},\sigma,\mathbf{s}\cdot c} + \frac{1}{2} D_{\mathbb{Z}^{nd},\sigma,-\mathbf{s}\cdot c}$. To prevent signatures from leaking the private key, we use rejection sampling that finds a positive integer M such that for all supports except a negligible fraction:

$$D_{\mathbb{Z}^{nd},\sigma} \leq M \cdot \left(\frac{1}{2} D_{\mathbb{Z}^{nd},\sigma,\mathbf{s}\cdot c} + \frac{1}{2} D_{\mathbb{Z}^{nd},\sigma,-\mathbf{s}\cdot c} \right)$$

It is thus sufficient to choose $M \geq \exp(\|\mathbf{s} \cdot c\|^2 / (2\sigma^2))$. Now we bound $\|\mathbf{s} \cdot c\|$. The random oracle H outputs a binary vector c with length n and weight κ (here we define $\mathcal{B}_\kappa$ to be the set of ternary vectors of length n with weight κ), and $\|\mathbf{s}\|_\infty$ is bounded by β, so $\|\mathbf{s} \cdot c\| \leq (\kappa \cdot \beta)\sqrt{nd}$. Hence, the number of repetitions M is approximately $\exp(\kappa^2 \beta^2 nd / (2\sigma^2))$.

Correctness. Let $(\mathbf{z}, c)$ be a valid signature for message μ. The rejection sampling shows that $\mathbf{z}$ follows a discrete Gaussian $D_{\mathbb{Z}^{nd},\sigma}$. By [29, Lemma 4.4], we have $\|\mathbf{z}\| \leq \eta \cdot \sigma\sqrt{nd}$, except with probability $\approx \eta^{nd} e^{nd/2(1-\eta^2)}$ for some small constant $\eta > 1$. In the security proof, we will also need $\|\mathbf{z}\|_\infty < q/4$. This is usually satisfied whenever $\|\mathbf{z}\| \leq \eta\sigma\sqrt{nd}$. Finally, check that $\langle \mathbf{a}, \mathbf{z} \rangle + q \cdot c = \langle \mathbf{a}, \mathbf{y} \rangle + (-1)^b \cdot c \cdot \langle \mathbf{a}, \mathbf{s} \rangle + q \cdot c \pmod{2q}$.

4.2 Security Proof

We sketch the proof that the signature in Figure 4 is secure under existential forgery using the Forking Lemma of Bellare-Neven [8] which follows similarly to [15]. We reduce the security of the signature to the variant MNTRU problem.

```
Algorithm $\mathsf{Gen}(q, n, d, \beta)$
Sample $\mathbf{f} = \{f_i\}_{i=1}^{d-1}$ and $f_d := 2g + 1$ where $f_i \hookleftarrow U_\beta^n$ and $g \hookleftarrow U_{\lfloor \beta/2 \rfloor}^n$
Compute $\mathbf{h} = (h_1, \ldots, h_{d-1}, -1)$ such that $\sum_{i=1}^{d-1} h_i f_i \equiv f_d \pmod q$
Set $\mathbf{a} = (2h_1, \cdots, 2h_{d-1}, q-2) \in R_{2q}^d$ and $\mathbf{s} = (f_1, \cdots, f_d) \in R_{2q}^d$
return $\mathsf{pk} := \mathbf{a}$ and $\mathsf{sk} := \mathbf{s}$
```

```
Algorithm $\mathsf{Sign}(\mathsf{sk}, \mu, \sigma)$ :
Sample $\mathbf{y} := (y_1, \ldots, y_d)$ where $y_i \hookleftarrow D_\sigma^n$
Compute $c = H(\langle \mathbf{a}, \mathbf{y} \rangle \pmod{2q}, \mu)$
Sample a random bit $b \in \{0, 1\}$
Compute $\mathbf{z} = (z_1, \ldots, z_d)$ where
    $z_i = y_i + (-1)^b \cdot c \cdot f_i$
return $(\mathbf{z}, c)$ with probability
    $1/\left(M \exp\left(-\frac{\|c \cdot \mathbf{s}\|^2}{2\sigma^2}\right) \cosh\left(\frac{\langle \mathbf{z}, c \cdot \mathbf{s} \rangle}{\sigma^2}\right)\right)$
```

```
Algorithm $\mathsf{Ver}(\mathsf{pk}, \mu, \mathbf{z}, c)$ :
if $\|\mathbf{z}\|_\infty < q/4$ and $\|\mathbf{z}\| < \eta\sigma\sqrt{nd}$ and
 $H(\langle \mathbf{a}, \mathbf{z} \rangle + q\,c \pmod{2q}, \mu) = c$ then
    return Accept
return $\perp$
```

Fig. 4. A BLISS-like signature scheme based on MNTRU

We construct two games, Hybrid 1 and Hybrid 2, as in Figure 5, and use them to simulate the genuine signature scheme. The distributions of outputs in Hybrid 1 and outputs in Hybrid 2 are the same due to rejection sampling. Thus, it is sufficient to show the genuine signature is statistically close to Hybrid 1.

Lemma 6. *Let $\mathcal{D}$ be an algorithm with the goal to distinguish the outputs of the genuine signing algorithm in Figure 4 and Hybrid 1 in Figure 5. Let $\mathcal{D}$ have access to two oracles: $\mathcal{O}_H$ and $\mathcal{O}_{\mathsf{Sign}}$. $\mathcal{O}_H$ is the hash oracle which, given an input x, outputs $H(x)$. $\mathcal{O}_{\mathsf{Sign}}$ is the oracle which, given an input, returns either the output of the signing algorithm or the output of Hybrid 1. If $\mathcal{D}$ makes at most q_H calls to $\mathcal{O}_H$ and q_S calls to $\mathcal{O}_{\mathsf{Sign}}$, then $Adv(\mathcal{D}) \leq q_S(q_H + q_S)2^{-n}$.*

We now prove the BLISS-like signature scheme in Figure 4 admits security against existential forgery under adaptive chosen-message attacks. First, we observe that if there exists an adversary capable of forging Hybrid 2 with advantage δ in polynomial time, then by the previous lemma, the adversary is capable of forging the genuine signature of Figure 4 with probability $\approx \delta$ in polynomial time. Thus, it is sufficient to reduce the variant MNTRU to the forging problem on Hybrid 2. We sketch it in the following theorem.

Theorem 1. *If there exists a polynomial-time algorithm $\mathcal{A}$ to forge the signature of Hybrid 2 with at most q_S signing queries to Hybrid 2 and q_H hash queries to the random oracle H, and it succeeds with probability δ, then there exists a polynomial-time algorithm that solves the* variant $\mathsf{MNTRU}_{q,n,d,D_1,D_2,B}$ *search*

Hybrid 1: $\mathsf{Sign}_1(\mathsf{sk}, \mu, \sigma)$

```
Sample $\mathbf{y} := (y_1, \ldots, y_d)$ where $y_i \hookleftarrow D_\sigma^n$
Sample $c \hookleftarrow \mathcal{B}_\kappa$
Sample a random bit $b$
Compute $\mathbf{z} = \mathbf{y} + (-1)^b \cdot c \cdot \mathbf{s}$
return $(\mathbf{z}, c)$ with probability
    $1/\left(M \exp\left(-\frac{\|c \cdot \mathbf{s}\|^2}{2\sigma^2}\right) \cosh\left(\frac{\langle \mathbf{z}, c \cdot \mathbf{s} \rangle}{\sigma^2}\right)\right)$
    Program $H(\langle \mathbf{a}, \mathbf{z} \rangle + q\,c \pmod{2q}, \mu) = c$
```

Hybrid 2: $\mathsf{Sign}_2(\sigma)$

```
Sample $c \hookleftarrow \mathcal{B}_\kappa$
Sample $\mathbf{z} = (z_1, \ldots, z_d)$ where $z_i \hookleftarrow D_\sigma^n$
return $(\mathbf{z}, c)$ with probability $1/M$
    Program $H(\langle \mathbf{a}, \mathbf{z} \rangle + q\,c \pmod{2q}, \mu) = c$
```

Fig. 5. Hybrid games of Figure 4

problem with advantage $\approx \delta^2/(q_S + q_H)$, where distributions D_1 and D_2 sample each coordinate-wise polynomial from $D_{\mathbb{Z},\sigma}^n$ and $B := 2\eta\sigma\sqrt{nd}$.

We sketch the proof of Lemma 6 and Theorem 1 in the full version of this work.

5 Security analysis and parameters

In this section, we discuss known attacks for the MNTRU assumptions based on lattice reduction [42,43] for MNTRU lattices. We assume that the variant iMNTRU problem used in our signatures admits a similar security of the same dimension. Let $\mathcal{N}$ be an $\mathsf{MNTRU}_{q,n,d,B}$ distribution, and a vector of polynomials $\mathbf{h} \in R_q^{d-1}$ be a sample from $\mathcal{N}$. The lattice associated to $\mathbf{h}$ is defined as

$$\Lambda_{\mathbf{h}} := \left\{(x_1, \ldots, x_d) \in R_q^d : x_1 h_1 + \ldots + x_{d-1} h_{d-1} + x_d = 0 \pmod{q}\right\}.$$

It has a basis generated by the columns of

$$\mathbf{B} := \begin{bmatrix} I_n & 0_n & \ldots & 0_n & 0_n \\ 0_n & I_n & \ldots & 0_n & 0_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_n & 0_n & \cdots & I_n & 0_n \\ -M_{h_1} & -M_{h_2} & \ldots & -M_{h_{d-1}} & qI_n \end{bmatrix}$$

The lattice $\mathcal{L}(\mathbf{B})$ has rank $d \times n$ and determinant q^n. Let $(\mathbf{f}, g)$ from $R_q^{d-1} \times R_q$ be a solution of a search $\mathsf{MNTRU}_{q,n,d,B}$ problem. One can verify that $(\mathbf{f}, g)$ is a short vector of $\Lambda_{\mathbf{h}}$ by the relation $\mathbf{B} \cdot \begin{bmatrix} v_{\mathbf{f}} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} v_{\mathbf{f}} \\ g \end{bmatrix}$. Thus if one can solve the SVP problem in $\Lambda_{\mathbf{h}}$, one can find a solution for the corresponding MNTRU problem.

We review the methodology for estimating the Core-SVP security in the full version of this work and use them to develop the concrete parameters in Table 1.

5.1 Concrete Instantiation

	I	II	III	IV	V	VI
Ring Dimension n	2048	1024	4096	2048	1283	2003
Module Rank d	2	4	2	3	3	2
Ring Modulus $\log_2(q)$	39.93	78.68	53.47	71.37	55.89	38.95
κ	32	37	28	32	35	32
r	21	22	34	33	18	20
γ	47668	80205	79918	91335	71583	48041
Acceptance Rate	0.237	0.238	0.238	0.238	0.202	0.233
Block-Size b	490	500	839	669	494	492
Public Key pk (bytes)	10272	10144	27680	18464	9013	9797
Signature Size $\mathbf{z}$ (bytes)	4400	6963	9262	9264	5824	4305

Table 1. Concrete parameters for signature in Section 3

We propose the concrete parameters for our signature scheme in Section 3, with an 128-bit security level achieved by using Theorem 3.1 of [26]. The size of the public key is $n \cdot \lceil \log q \rceil + 256$ bits when using a 256-bit seed to generate the randomness. The signature size is $n \cdot (d-1) \cdot \lceil \log 2(\gamma - \beta \cdot \kappa) \rceil + \kappa(\log(n) + 1)$ bits. For all parameters, the rejection rate is chosen such that the repetition rate is approximately 4.2–4.3, which is comparable to the rejection rate of the 127 bit security scheme in [26] which has the smallest signature size for schemes provable in the QROM. The secret key is taken to be ternary in all cases, that is to say that $\beta = 1$ in all columns in the table. Columns I-IV are arranged with increasing signature size. These four columns are proven secure in Section 3 of this work. Columns I and II have BKZ block sizes close to the bound of 128 bit security while columns III and IV have block sizes suitable for higher security considerations. Note that columns II and IV have very large prime moduli, making them potentially weak to subfield attacks [3,27]. To heuristically combat this, one may change β to increase the space of valid secret keys at the cost of signature and public key sizes. Updated choices for β resilient to subfield attacks are left to future works. The optimal provably secure signature size in [26] is 5690 bytes and has public key size 7712 bytes. Comparing this to column I in the table we see that our scheme achieves comparable security and acceptance rates with a signature 77% the size of theirs at the expense of having public key 133% the size. This tradeoff makes their scheme have better overall channel weight if one message is to be signed, but if more than one is to be sent, then our parameter set in column I has a lower overall channel weight.

Columns V and VI use the NTRU-prime [9] like polynomials with irreducible polynomials $x^n - x - 1$ for prime n; thus the underlying rings do not correspond to power-of-two cyclotomics. The flexibility of choosing n leaves room for improvement on provable parameters, as one sees that NTRU-prime constructions give the smallest signature size (VI) and smallest public key size (V). We remark that the security of these two columns is not proven here since our proofs (e.g. Lemma 3) use the underlying ring structure. We leave them to future works.

For the BLISS-like signature scheme in Section 4, the public key and the secret key are vectors of polynomials in $U_{R_q}^{d-1}$ and U_{β}^{nd}, thus amounting to $n \cdot (d-1) \cdot \lceil \log q \rceil$ bits and $n \cdot d \cdot \lceil \log 2\beta \rceil$ bits, respectively. The signature is $(\mathbf{z}, c)$, where $\mathbf{z} \in R_q^d$ with $\|\mathbf{z}\|_\infty < q/4$, and c sampled from the set of binary vectors of length n with Hamming weight κ. Thereby, the size of signature is $(n \cdot d \cdot \lceil \log(q/4) \rceil + n)$ bits. The signature in Section 4 utilizes the same framework as the BLISS signature. We expect it yields more flexibility in selecting parameters due to the usage of module lattices. It remains an interesting question to understand whether the BLISS-like signature is secure in the QROM, and thus we leave the parameter selection for future work.

Acknowledgement

The authors thank the reviewers for their helpful discussions and remarks.

References

1. M. Abdalla, P.-A. Fouque, V. Lyubashevsky, and M. Tibouchi. Tightly-secure signatures from lossy identification schemes. In D. Pointcheval and T. Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 572–590. Springer, Heidelberg, Apr. 2012.
2. M. Ajtai. Generating hard instances of lattice problems (extended abstract). In *28th ACM STOC*, pages 99–108. ACM Press, May 1996.
3. M. R. Albrecht, S. Bai, and L. Ducas. A subfield lattice attack on overstretched NTRU assumptions - cryptanalysis of some FHE and graded encoding schemes. In M. Robshaw and J. Katz, editors, *CRYPTO 2016, Part I*, volume 9814 of *LNCS*, pages 153–178. Springer, Heidelberg, Aug. 2016.
4. E. Alkim, P. S. L. M. Barreto, N. Bindel, J. Krämer, P. Longa, and J. E. Ricardini. The lattice-based digital signature scheme qTESLA. In M. Conti, J. Zhou, E. Casalicchio, and A. Spognardi, editors, *ACNS 20, Part I*, volume 12146 of *LNCS*, pages 441–460. Springer, Heidelberg, Oct. 2020.
5. E. Alkim, N. Bindel, J. A. Buchmann, Ö. Dagdelen, E. Eaton, G. Gutoski, J. Krämer, and F. Pawlega. Revisiting TESLA in the quantum random oracle model. In T. Lange and T. Takagi, editors, *Post-Quantum Cryptography - 8th International Workshop, PQCrypto 2017*, pages 143–162. Springer, Heidelberg, 2017.
6. J. Alwen and C. Peikert. Generating shorter bases for hard random lattices. *Theor. Comp. Sys.*, 48(3):535–553, Apr. 2011.

7. S. Bai and S. D. Galbraith. An improved compression technique for signatures based on learning with errors. In J. Benaloh, editor, *CT-RSA 2014*, volume 8366 of *LNCS*, pages 28–47. Springer, Heidelberg, Feb. 2014.
8. M. Bellare and G. Neven. Multi-signatures in the plain public-key model and a general forking lemma. In A. Juels, R. N. Wright, and S. De Capitani di Vimercati, editors, *ACM CCS 2006*, pages 390–399. ACM Press, Oct. / Nov. 2006.
9. D. J. Bernstein, B. B. Brumley, M.-S. Chen, C. Chuengsatiansup, T. Lange, A. Marotzke, B.-Y. Peng, N. Tuveri, C. van Vredendaal, and B.-Y. Yang. NTRU Prime. Technical report, National Institute of Standards and Technology, 2020. available at `https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions`.
10. N. Bindel, S. Akleylek, E. Alkim, P. S. L. M. Barreto, J. Buchmann, E. Eaton, G. Gutoski, J. Kramer, P. Longa, H. Polat, J. E. Ricardini, and G. Zanon. qTESLA. Technical report, National Institute of Standards and Technology, 2019. available at `https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions`.
11. Z. Brakerski, A. Langlois, C. Peikert, O. Regev, and D. Stehlé. Classical hardness of learning with errors. In D. Boneh, T. Roughgarden, and J. Feigenbaum, editors, *45th ACM STOC*, pages 575–584. ACM Press, June 2013.
12. Y. Chen, N. Genise, and P. Mukherjee. Approximate trapdoors for lattices and smaller hash-and-sign signatures. In S. D. Galbraith and S. Moriai, editors, *ASIACRYPT 2019, Part III*, volume 11923 of *LNCS*, pages 3–32. Springer, Heidelberg, Dec. 2019.
13. J. H. Cheon, D. Kim, T. Kim, and Y. Son. A new trapdoor over module-NTRU lattice and its application to ID-based encryption. Cryptology ePrint Archive, Report 2019/1468, 2019. `https://eprint.iacr.org/2019/1468`.
14. C. Chuengsatiansup, T. Prest, D. Stehlé, A. Wallet, and K. Xagawa. ModFalcon: Compact signatures based on module-NTRU lattices. In H.-M. Sun, S.-P. Shieh, G. Gu, and G. Ateniese, editors, *ASIACCS 20*, pages 853–866. ACM Press, Oct. 2020.
15. L. Ducas, A. Durmus, T. Lepoint, and V. Lyubashevsky. Lattice signatures and bimodal Gaussians. In R. Canetti and J. A. Garay, editors, *CRYPTO 2013, Part I*, volume 8042 of *LNCS*, pages 40–56. Springer, Heidelberg, Aug. 2013.
16. L. Ducas, V. Lyubashevsky, and T. Prest. Efficient identity-based encryption over NTRU lattices. In P. Sarkar and T. Iwata, editors, *ASIACRYPT 2014, Part II*, volume 8874 of *LNCS*, pages 22–41. Springer, Heidelberg, Dec. 2014.
17. A. Fiat and A. Shamir. How to prove yourself: Practical solutions to identification and signature problems. In A. M. Odlyzko, editor, *CRYPTO'86*, volume 263 of *LNCS*, pages 186–194. Springer, Heidelberg, Aug. 1987.
18. P.-A. Fouque, J. Hoffstein, P. Kirchner, V. Lyubashevsky, T. Pornin, T. Prest, T. Ricosset, G. Seiler, W. Whyte, and Z. Zhang. FALCON: Fast-Fourier Lattice-based Compact Signatures over NTRU. `https://falcon-sign.info/`, 2017.
19. N. Genise, C. Gentry, S. Halevi, B. Li, and D. Micciancio. Homomorphic encryption for finite automata. In S. D. Galbraith and S. Moriai, editors, *ASIACRYPT 2019, Part II*, volume 11922 of *LNCS*, pages 473–502. Springer, Heidelberg, Dec. 2019.
20. C. Gentry, C. Peikert, and V. Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In R. E. Ladner and C. Dwork, editors, *40th ACM STOC*, pages 197–206. ACM Press, May 2008.
21. T. Güneysu, V. Lyubashevsky, and T. Pöppelmann. Practical lattice-based cryptography: A signature scheme for embedded systems. In E. Prouff and P. Schaumont,

editors, *CHES 2012*, volume 7428 of *LNCS*, pages 530–547. Springer, Heidelberg, Sept. 2012.

22. J. Hoffstein, N. Howgrave-Graham, J. Pipher, J. H. Silverman, and W. Whyte. NTRUSIGN: Digital signatures using the NTRU lattice. In M. Joye, editor, *CT-RSA 2003*, volume 2612 of *LNCS*, pages 122–140. Springer, Heidelberg, Apr. 2003.
23. J. Hoffstein, J. Pipher, and J. H. Silverman. NTRU: A new high speed public key cryptosystem, 1996. Draft Distributed at Crypto'96, available at `http://web.securityinnovation.com/hubfs/files/ntru-orig.pdf`.
24. J. Hoffstein, J. Pipher, and J. H. Silverman. Ntru: A ring-based public key cryptosystem. In J. P. Buhler, editor, *Algorithmic Number Theory*, pages 267–288, Berlin, Heidelberg, 1998. Springer Berlin Heidelberg.
25. J. Hoffstein, J. Pipher, and J. H. Silverman. NTRU: A ring-based public key cryptosystem. In *ANTS*, pages 267–288, 1998.
26. E. Kiltz, V. Lyubashevsky, and C. Schaffner. A concrete treatment of Fiat-Shamir signatures in the quantum random-oracle model. In J. B. Nielsen and V. Rijmen, editors, *EUROCRYPT 2018, Part III*, volume 10822 of *LNCS*, pages 552–586. Springer, Heidelberg, Apr. / May 2018.
27. P. Kirchner and P.-A. Fouque. Revisiting lattice attacks on overstretched NTRU parameters. In J.-S. Coron and J. B. Nielsen, editors, *EUROCRYPT 2017, Part I*, volume 10210 of *LNCS*, pages 3–26. Springer, Heidelberg, Apr. / May 2017.
28. V. Lyubashevsky. Fiat-Shamir with aborts: Applications to lattice and factoring-based signatures. In M. Matsui, editor, *ASIACRYPT 2009*, volume 5912 of *LNCS*, pages 598–616. Springer, Heidelberg, Dec. 2009.
29. V. Lyubashevsky. Lattice signatures without trapdoors. In D. Pointcheval and T. Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 738–755. Springer, Heidelberg, Apr. 2012.
30. V. Lyubashevsky, L. Ducas, E. Kiltz, T. Lepoint, P. Schwabe, G. Seiler, D. Stehlé, and S. Bai. CRYSTALS-DILITHIUM. Technical report, National Institute of Standards and Technology, 2020. available at `https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions`.
31. V. Lyubashevsky and D. Micciancio. Generalized compact Knapsacks are collision resistant. In M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener, editors, *ICALP 2006, Part II*, volume 4052 of *LNCS*, pages 144–155. Springer, Heidelberg, July 2006.
32. V. Lyubashevsky, C. Peikert, and O. Regev. On ideal lattices and learning with errors over rings. In H. Gilbert, editor, *EUROCRYPT 2010*, volume 6110 of *LNCS*, pages 1–23. Springer, Heidelberg, May / June 2010.
33. D. Micciancio. Generalized compact knapsacks, cyclic lattices, and efficient one-way functions from worst-case complexity assumptions. In *43rd FOCS*, pages 356–365. IEEE Computer Society Press, Nov. 2002.
34. D. Micciancio and C. Peikert. Trapdoors for lattices: Simpler, tighter, faster, smaller. In D. Pointcheval and T. Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 700–718. Springer, Heidelberg, Apr. 2012.
35. D. Micciancio and O. Regev. Worst-case to average-case reductions based on Gaussian measures. In *45th FOCS*, pages 372–381. IEEE Computer Society Press, Oct. 2004.
36. C. Peikert. A decade of lattice cryptography. Found. Trends Theor. Comput. Sci., 10(4)., 2016. `http://eprint.iacr.org/`.
37. C. Peikert and A. Rosen. Lattices that admit logarithmic worst-case to average-case connection factors. In D. S. Johnson and U. Feige, editors, *39th ACM STOC*, pages 478–487. ACM Press, June 2007.

38. A. Pellet-Mary and D. Stehlé. On the hardness of the ntru problem. Cryptology ePrint Archive, Report 2021/821, 2021. https://ia.cr/2021/821.
39. T. Prest, P.-A. Fouque, J. Hoffstein, P. Kirchner, V. Lyubashevsky, T. Pornin, T. Ricosset, G. Seiler, W. Whyte, and Z. Zhang. FALCON. Technical report, National Institute of Standards and Technology, 2020. available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.
40. O. Regev. On lattices, learning with errors, random linear codes, and cryptography. In H. N. Gabow and R. Fagin, editors, *37th ACM STOC*, pages 84–93. ACM Press, May 2005.
41. O. Regev. Lattice-based cryptography (invited talk). In C. Dwork, editor, *CRYPTO 2006*, volume 4117 of *LNCS*, pages 131–141. Springer, Heidelberg, Aug. 2006.
42. C.-P. Schnorr. A hierarchy of polynomial time lattice basis reduction algorithms. *Theor. Comput. Sci.*, 53:201–224, 1987.
43. C.-P. Schnorr and M. Euchner. Lattice basis reduction: Improved practical algorithms and solving subset sum problems. *Mathmatical Programming*, 66:181–199, 1994.
44. D. Stehlé, R. Steinfeld, K. Tanaka, and K. Xagawa. Efficient public key encryption based on ideal lattices. In M. Matsui, editor, *ASIACRYPT 2009*, volume 5912 of *LNCS*, pages 617–635. Springer, Heidelberg, Dec. 2009.
45. Z. Zhang, C. Chen, J. Hoffstein, and W. Whyte. pqNTRUSign. Technical report, National Institute of Standards and Technology, 2017. available at https://csrc.nist.gov/projects/post-quantum-cryptography/round-1-submissions.