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) [2,40], the Learning With Errors problem (LWE) [45,46,37,14] and the NTRU problem [28,26].

* 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 [2,24,33]. The LWE problem, introduced by Regev [45,46], is extensively used as a security foundation for encryption, signatures and many others [46,24,18,33]. For efficiency, many practical lattice-based cryptosystems are based on assumptions on structured lattices such as the Ring-LWE [37,50], Ring-SIS[38,35,42] and the NTRU problems [27,29]. Introduced by Hoffstein, Pipher and Silverman [27,29], 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 [27,29], the NTRU assumption has been used extensively in various cryptographic constructions such as encryption, signature and many others [26,18,19]. Little is known on the complexity reduction aspects of the NTRU problem (see also [43,41] 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 [24,33,18,21,13]. 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 [24] (GPV). The second framework, proposed by Lyubashevsky [32,33], utilizes the Fiat-Shamir [20] 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 [24], 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 [8,39]. 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 [19], 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 [16,17] 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 [16,17] 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 [44], pqNTRUSign [51], 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 [20]. In [32,33], 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 [18], 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 [18] also devised an efficient signature scheme using the NTRU assumption. Follow-up work such as [25,18,9,6] 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 [10] 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 [6,5]. 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 [30] 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 [6,5]. A concrete instantiation of [30] is to adapt and to modify the Dilithium signature scheme [34], which has tight secure reductions from Module-SIS (MSIS) and Module-LWE (MLWE). A concrete instantiation of the techniques in [6,5] is given in the qTESLA signature [13], 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 [30] 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 [30], starting from an identification scheme and applying the Fiat-Shamir transform. The second scheme is analogous to the BLISS [18] 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 [30] and BLISS [18], instantiated with the (inhomogeneous) Module-NTRU assumptions.

We first generalize the Module-NTRU problem proposed in [16,17] 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 [30] 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 [18], 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 [23], 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 [15] 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 [39]. 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 [34] and qTESLA [13] 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 [34] and qTESLA [13] are both built on the Module-LWE assumptions.

Finally, compared to the BLISS signature [18], 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 [43] 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 $\mathbf{v^t}$ 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 give background on lattices in Appendix A. We will use the rejection sampling lemma from [33] to ensure the output signature does not leak information about the secret key. We review the definition of various distributions and rejection sampling lemma in Appendix B. We also review the background on identification, digital signatures and the Fiat-Shamir transform in Appendices C, D and E.

2.1 (Inhomogeneous) Module-NTRU

As a generalization of NTRU, the Module-NTRU (MNTRU) problem was introduced in [16,17], 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 [16,17].

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 [33]. 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 [1,5,30] 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 [18] for efficient instantiation of random oracles.

3.1 A lossy identification scheme

As in [1,30], 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 [16,17], 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 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

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

1 : Sample $\mathbf{f} = \{f_i\}_{i=1}^{d}$ and t, where $f_i \hookleftarrow U_\beta^n$ and $t \hookleftarrow U_{R_q}$

2 : Compute $\mathbf{h} = (h_1, \ldots, h_{d-1}, 1)$ such that $\sum_{i=1}^{d} h_i f_i \equiv t \pmod q$

3 : **return** $\mathsf{pk} := (\mathbf{h}, t)$ and $\mathsf{sk} := \mathbf{f}$

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

4 : Sample $\mathbf{y} = \{y_i\}_{i=1}^{d-1}$ where $y_i \hookleftarrow U_\gamma^n$

5 : Compute $u = \left\lfloor \sum_{i=1}^{d-1} h_i y_i \pmod q \right\rceil_r$

6 : **return** u

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

7 : Compute $\mathbf{z} := (z_1, \ldots, z_{d-1})$ where $z_i = y_i + c \cdot f_i$

8 : Compute $w = \sum_{i=1}^{d-1} h_i y_i - c \cdot f_d \pmod q$

9 : **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**

10 : **return** $\perp$

11 : **return** $\mathbf{z}$

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

12 : **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**

13 : **return** Accept

14 : **return** Reject

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

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 [30]. For Lemmas 1 to 5, we state them and sketch the proofs in Appendix F.

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*

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

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 $u = \left\lfloor \sum_{i=1}^{d-1} h_i z_i - t \cdot c \right\rceil_r$.

Algorithm $\mathsf{Trans}(\mathsf{sk})$

1 : Sample $\mathbf{y} = \{y_i\}_{i=1}^{d-1}$ where $y_i \hookleftarrow U_\gamma^n$

2 : Compute $u = \left\lfloor \sum_{i=1}^{d-1} h_i y_i \pmod q \right\rceil_r$

3 : Sample $c \hookleftarrow \mathcal{B}_\kappa$

4 : Compute $\mathbf{z} = \{z_i\}_{i=1}^{d-1}$ where $z_i = y_i + c \cdot f_i$

5 : Compute $w = \sum_{i=1}^{d-1} h_i y_i - c \cdot f_d \pmod q$

6 : **if** any $\|z_i\|_\infty > \gamma - B$ **return** $\perp$

7 : **if** $\left\|[w]_r\right\|_\infty \geq 2^{r-1} - B$
or $\|w\|_\infty \geq \lfloor q/2 \rfloor - B$ **return** $\perp$

8 : **return** $(\mathbf{z}, c)$

Algorithm $\mathsf{Sim}(\mathsf{pk})$

9 : With probability $1 - \left(\frac{|U_{\gamma-B}|}{|U_\gamma|}\right)^{n(d-1)}$
return $\perp$

10 : Sample $\mathbf{z} = \{z_i\}_{i=1}^{d-1}$ where $z_i \hookleftarrow U_{\gamma-B}^n$

11 : Sample $c \hookleftarrow \mathcal{B}_\kappa$

12 : Compute $w' = \sum_{i=1}^{d-1} h_i z_i - t \cdot c \pmod q$

13 : **if** $\left\|\lceil w' \rfloor_r\right\|_\infty \geq 2^{r-1} - B$
or $\left\|w'\right\|_\infty \geq \lfloor q/2 \rfloor - B$ **return** $\perp$

14 : **return** $(\mathbf{z}, c)$

Fig. 2. Transcript algorithm and simulation algorithm

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

1 : Sample $\mathbf{h} = (h_1, \ldots, h_{d-1}, 1)$ and t, where $h_i \hookleftarrow U_{R_q}$ and $t \hookleftarrow U_{R_q}$

2 : **return** $\mathsf{pk} := (\mathbf{h}, t)$

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

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 in Figure 7 of Appendix F. Theorem 3.1 of [30] (also see Appendix E) 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 [18], 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 $D^n_{\mathbb{Z},\sigma}$. 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 [18] to compress the signature.

Rejection Sampling. The rejection sampling follows the same as [18]. 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 [33, 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 [10] which follows similarly to [18]. We reduce the security of the signature to the variant MNTRU problem.

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.

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

1 : Sample $\mathbf{f} = \{f_i\}_{i=1}^{d-1}$ and $f_d := 2g + 1$ where $f_i \hookleftarrow U_\beta^n$ and $g \hookleftarrow U_{\lfloor \beta/2 \rfloor}^n$

2 : Compute $\mathbf{h} = (h_1, \ldots, h_{d-1}, -1)$ such that $\sum_{i=1}^{d-1} h_i f_i \equiv f_d \pmod q$

3 : Set $\mathbf{a} = (2h_1, \cdots, 2h_{d-1}, q-2) \in R_{2q}^d$ and $\mathbf{s} = (f_1, \cdots, f_d) \in R_{2q}^d$

4 : **return** $\mathsf{pk} := \mathbf{a}$ and $\mathsf{sk} := \mathbf{s}$

Algorithm $\mathsf{Sign}(\mathsf{sk}, \mu, \sigma)$:

5 : Sample $\mathbf{y} := (y_1, \ldots, y_d)$ where $y_i \hookleftarrow D_\sigma^n$

6 : Compute $c = H(\langle \mathbf{a}, \mathbf{y} \rangle \pmod{2q}, \mu)$

7 : Sample a random bit $b \in \{0, 1\}$

8 : Compute $\mathbf{z} = (z_1, \ldots, z_d)$ where $z_i = y_i + (-1)^b \cdot c \cdot f_i$

9 : **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 $\mathsf{Ver}(\mathsf{pk}, \mu, \mathbf{z}, c)$:

10 : **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**

11 : **return** Accept

12 : **return** $\perp$

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

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 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 Appendix G.

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

1 : Sample $\mathbf{y} := (y_1, \ldots, y_d)$ where $y_i \hookleftarrow D_\sigma^n$
2 : Sample $c \hookleftarrow \mathcal{B}_\kappa$
3 : Sample a random bit b
4 : Compute $\mathbf{z} = \mathbf{y} + (-1)^b \cdot c \cdot \mathbf{s}$
5 : **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: $\mathsf{Sign}_2(\sigma)$

1 : Sample $c \hookleftarrow \mathcal{B}_\kappa$
2 : Sample $\mathbf{z} = (z_1, \ldots, z_d)$ where $z_i \hookleftarrow D_\sigma^n$
3 : **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

5 Security analysis and parameters

In this section, we discuss known attacks for the MNTRU assumptions based on lattice reduction [47,49] 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 Appendix H 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 [30]. 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 [30] 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 [3,31]. 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 [30] 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 [12] 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. M. R. Albrecht, F. Göpfert, F. Virdia, and T. Wunderer. Revisiting the expected cost of solving uSVP and applications to LWE. In T. Takagi and T. Peyrin, editors, *ASIACRYPT 2017, Part I*, volume 10624 of *LNCS*, pages 297–322. Springer, Heidelberg, Dec. 2017.
5. 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.
6. 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.
7. E. Alkim, L. Ducas, T. Pöppelmann, and P. Schwabe. Post-quantum key exchange - A new hope. In T. Holz and S. Savage, editors, *USENIX Security 2016*, pages 327–343. USENIX Association, Aug. 2016.
8. J. Alwen and C. Peikert. Generating shorter bases for hard random lattices. *Theor. Comp. Sys.*, 48(3):535–553, Apr. 2011.

9. 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.
10. 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.
11. M. Bellare, B. Poettering, and D. Stebila. From identification to signatures, tightly: A framework and generic transforms. In J. H. Cheon and T. Takagi, editors, *ASIACRYPT 2016, Part II*, volume 10032 of *LNCS*, pages 435–464. Springer, Heidelberg, Dec. 2016.
12. 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`.
13. 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`.
14. 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.
15. 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.
16. 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`.
17. 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.
18. 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.
19. 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.
20. 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.
21. 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.
22. N. Gama and P. Q. Nguyen. Predicting lattice reduction. In N. P. Smart, editor, *EUROCRYPT 2008*, volume 4965 of *LNCS*, pages 31–51. Springer, Heidelberg, Apr. 2008.

23. 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.
24. 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.
25. 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.
26. 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.
27. 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`.
28. 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.
29. J. Hoffstein, J. Pipher, and J. H. Silverman. NTRU: A ring-based public key cryptosystem. In *ANTS*, pages 267–288, 1998.
30. 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.
31. 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.
32. 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.
33. 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.
34. 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`.
35. 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.
36. V. Lyubashevsky and G. Neven. One-shot verifiable encryption from lattices. Cryptology ePrint Archive, Report 2017/122, 2017. `https://eprint.iacr.org/2017/122`.
37. 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.
38. 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.

39. 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.
40. 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.
41. C. Peikert. A decade of lattice cryptography. Found. Trends Theor. Comput. Sci., 10(4)., 2016. `http://eprint.iacr.org/`.
42. 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.
43. 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`.
44. 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`.
45. 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.
46. O. Regev. Lattice-based cryptography (invited talk). In C. Dwork, editor, *CRYPTO 2006*, volume 4117 of *LNCS*, pages 131–141. Springer, Heidelberg, Aug. 2006.
47. C.-P. Schnorr. A hierarchy of polynomial time lattice basis reduction algorithms. *Theor. Comput. Sci.*, 53:201–224, 1987.
48. C.-P. Schnorr. Lattice reduction by random sampling and birthday methods. In *STACS*, pages 145–156, 2003.
49. C.-P. Schnorr and M. Euchner. Lattice basis reduction: Improved practical algorithms and solving subset sum problems. *Mathmatical Programming*, 66:181–199, 1994.
50. 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.
51. 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`.

A Euclidean lattices

Let $\mathbf{B} \in \mathbb{Q}^{m \times n}$ be a matrix of rank n. The lattice $\mathcal{L}$ generated by $\mathbf{B}$ is defined as $\mathcal{L}(\mathbf{B}) = \{\mathbf{Bx} \mid \forall \mathbf{x} \in \mathbb{Z}^n\}$, and the matrix $\mathbf{B}$ is called a basis of $\mathcal{L}(\mathbf{B})$ (or just $\mathcal{L}$). We let $\mathbf{B}^* = (\mathbf{b}_1^*, \cdots, \mathbf{b}_n^*)$ denote the Gram–Schmidt orthogonalization of $\mathbf{B}$. The determinant of a lattice $\mathcal{L}(\mathbf{B})$ is defined as $\mathrm{Vol}(\mathcal{L}(\mathbf{B})) = \prod_{i \leq n} \|\mathbf{b}_i^*\|$.

The ℓ_2-norm of a shortest non-zero vector in a lattice $\mathcal{L}$ is denoted by $\lambda_1(\mathcal{L})$, which is called the minimum of $\mathcal{L}$. This can be extended successively: For any lattice $\mathcal{L}$, the i-th minimum $\lambda_i(\mathcal{L})$ is the radius of the smallest ball with center at the origin and containing i linearly independent lattice vectors. With $\mathcal{B}(\mathbf{0}, r)$ denoting a ball at origin of radius r, we have $\lambda_i(\mathcal{L}) = \inf\{r : \dim(\mathrm{span}(\mathcal{L} \cap \mathcal{B}(\mathbf{0}, r))) \geq i\}$.

Minkowski's Convex Body Theorem states that $\lambda_1(\mathcal{L}) \leq 2 \cdot v_n^{-1/n} \cdot \mathrm{Vol}(\mathcal{L})^{1/n}$, where v_n is the volume of an n-dimensional Euclidean ball of radius 1. The average version of the Minkowski's Theorem is often known as the Gaussian heuristic: the λ_1 of a random n-dimensional lattice is asymptotically

$$\mathrm{GH}(\mathcal{L}) = v_n^{-1/n} \cdot \mathrm{Vol}(\mathcal{L})^{1/n}. \tag{1}$$

The quality of a full-rank lattice $\mathcal{L}$ of rank k is measured by the *root Hermite factor* δ so that $\|\mathbf{b}_1\| = \delta^k \mathrm{Vol}(\mathcal{L})^{1/k}$. For $i \leq n$, we let $\pi_i(\mathbf{v})$ denote the orthogonal projection of $\mathbf{v}$ onto the linear subspace $(\mathbf{b}_1, \cdots, \mathbf{b}_{i-1})^{\perp}$. For $i < j \leq n$, we let $B_{[i,j]}$ denote the local block $(\pi_i(\mathbf{b}_i), \cdots, \pi_i(\mathbf{b}_j))$, and $\mathcal{L}_{[i,j]}$ denote the lattice generated by $B_{[i,j]}$.

B Distribution and Rejection sampling

Let D be a probability distribution. We let $\mathrm{Supp}(D) = \{x : D(x) \neq 0\}$ denote its support. For a finite set X, we use U_X to denote the uniform distribution over X - e.g., U_{R_q} denotes the uniform distribution on $\mathbb{Z}_q[x]/(x^n + 1)$. For a positive number β, the uniform distribution on $[-\beta, \beta]$ is denoted as U_β. For a distribution D, we write $x \hookleftarrow D$ to denote that x is sampled from D. We denote $D(x)$ the probability of $x \leftarrow D$. A polynomial f of degree $(n-1)$ is identified as its coefficient vector and $f \hookleftarrow D^n$ denotes each coordinate of f is sampled from D independently. The statistical distance between two distributions D_1 and D_2 over a countable support X is $\Delta(D_1, D_2) = \frac{1}{2} \sum_{x \in X} |D_1(x) - D_2(x)|$. This definition is extended in the natural way to continuous distributions. If $f : X \to \mathbb{R}$ takes non-negative values, then for all countable $Y \subseteq X$, we define $f(Y) = \sum_{t \in Y} f(t)$. A function $f(\lambda)$ is negligible if it is $\lambda^{-\omega(1)}$, where λ often denotes the security parameter in our context. A probability density function $p(\lambda)$ is overwhelming if it is $1 - \lambda^{-\omega(1)}$. The distinguishing advantage of an algorithm $\mathcal{A}$ between two distributions D_0 and D_1 is defined as $\mathrm{Adv}_{\mathcal{A}}(D_0, D_1) = |\Pr_{x \hookleftarrow D_0}[\mathcal{A}(x) = 1] - \Pr_{x \hookleftarrow D_1}[\mathcal{A}(x) = 1]|$, where the probabilities are taken over the randomness of the input x and the internal randomness of $\mathcal{A}$. Algorithm $\mathcal{A}$ is called an (ε, T)-distinguisher if it runs in time $\leq T$ and if $\mathrm{Adv}_{\mathcal{A}}(D_0, D_1) \geq \varepsilon$.

A distribution D defined on $\mathbb{R}$ is B-bounded for some positive real B, if the probability that $x \hookleftarrow D$ has absolute value greater than B is negligible w.r.t. the system parameter, which will be clear from the context. In the case where D is over $\mathbb{Z}_q$, we will assume implicitly that $B \leq (q-1)/2$. A B-bounded distribution D is said to be balanced if $\Pr[D \leq 0] \geq 1/2$ and $\Pr[D \geq 0] \geq 1/2$. The notion of B-bounded distribution extends coordinate-wise if the support of D is $\mathbb{R}^n$ and $\mathbb{R}^{n \times n}$. For example, we will consider the case when the polynomials f are sampled from a B-bounded distribution, thus each coefficient of f sampled is B-bounded.

For a vector $\mathbf{c} \in \mathbb{R}^n$ and a real $s > 0$, the Gaussian function $\rho_{s,\mathbf{c}}$ with standard deviation s and center $\mathbf{c}$ is defined as $\rho_{s,\mathbf{c}}(\mathbf{x}) = \exp(-\pi\|\mathbf{x}-\mathbf{c}\|^2/s^2), \forall \mathbf{x} \in \mathbb{R}^n$. The Gaussian distribution function is $D_{s,\mathbf{c}}(\mathbf{x}) = \rho_{s,\mathbf{c}}(\mathbf{x})/s^n$. When $\mathbf{c} = \mathbf{0}$, we omit the subscript $\mathbf{c}$. The discrete Gaussian distribution over a lattice $\Lambda \subseteq \mathbb{R}^n$, with standard deviation $s > 0$ and center $\mathbf{c}$ is defined as: $D_{\Lambda,s,\mathbf{c}} = \rho_{s,\mathbf{c}}(\mathbf{x})/\rho_{s,\mathbf{c}}(\Lambda), \forall \mathbf{x} \in \Lambda$. When the center is $\mathbf{0}$, we may omit the subscript $\mathbf{c}$. When the lattice $\Lambda = \mathbb{Z}$, we may also omit the subscript Λ.

We use the following rejection sampling lemma from [33] to ensure the output signature is not leaking information about the secret key.

Lemma 7. *Let V be an arbitrary set, and $h : V \to \mathbb{R}$ and $f : \mathbb{Z}_m \to \mathbb{R}$ be probability distributions. If $g_v : \mathbb{Z}_m \to \mathbb{R}$ is a family of probability distributions indexed by all $v \in V$ with the property that*

$$\exists M \in \mathbb{R} \text{ such that } \forall v, \Pr\left[Mg_v(z) \geq f(z); z \leftarrow f\right] \geq 1 - \epsilon$$

then the distribution of the output of the following algorithm is $\frac{1-\epsilon}{M}$ indistinguishable:

1. *$v \leftarrow h, z \leftarrow g_v$, return (z, v) with probability $\min\left(\frac{f(z)}{Mg_v(z)}, 1\right)$.*
2. *$v \leftarrow h, z \leftarrow f$, return (z, v) with probability $\frac{1}{M}$.*

C Identification and signature schemes

We recall some basic notion on the canonical identification schemes, following [30]. A canonical identification scheme is a three-move protocol between two parties: a prover P and a verifier V. In the three-move protocol, the prover sends a commitment W to the verifier, then the verifier selects a random challenge c and sends it to P. Upon receiving c, the prover sends back a response Z to the verifier. In the end, the verifier makes a deterministic decision about the received Z.

Definition 5 (Canonical identification scheme). *A canonical identification scheme is a tuple of classical* ppt *algorithms* $\mathsf{ID} := (\mathsf{IGen}, \mathsf{P}, \mathsf{V})$.

- *The key generation algorithm* IGen *takes as input a security parameter λ and returns the public and secret keys* $(\mathsf{pk}, \mathsf{sk})$. *The public key defines the set of challenges* ChSet, *the set of commitments* WSet, *and the set of responses* ZSet.

- *The prover algorithm* P *consists of two sub-algorithms:* P_1 *takes as input the secret key* sk *and returns a commitment* $W \in \mathsf{WSet}$ *and a state* St*;* P_2 *takes as inputs the secret key* sk*, a commitment* W*, a challenge* c*, and a state* St *and returns a response* $Z \in \mathsf{ZSet} \cup \{\perp\}$*, where* $\perp \notin \mathsf{ZSet}$ *is a special symbol indicating failure.*
- *The verifier algorithm* V *takes as inputs the public key* pk *and the transcript* (W, c, Z) *and outputs* 1 *(acceptance) or* 0 *(rejection).*

If $Z = \perp$, then we set $(W, c, Z) = (\perp, \perp, \perp)$. The triple $(W, c, Z) \in \mathsf{WSet} \times \mathsf{ChSet} \times \mathsf{ZSet} \cup \{(\perp, \perp, \perp)\}$ generated in this way is called a *transcript* and denoted as $\mathsf{Trans}(\mathsf{sk})$. Given the public key pk, a transcript is valid if $V(\mathsf{pk}, W, c, Z) = 1$.

We say that ID has *correctness error* δ if, for all public and secret keys generated by IGen all possible transcripts in $\mathsf{WSet} \times \mathsf{ChSet} \times \mathsf{ZSet}$ with $Z \neq \perp$ are valid, and the probability that a honestly generated transcript is $(\perp, \perp, \perp)$ is bounded above by δ. If the most likely probability of a random variable W that is chosen from a discrete distribution D is $2^{-\alpha}$, then we write $H_\infty(W|W \hookleftarrow D) = \alpha$. We say a canonical identification scheme ID has α bits of *min-entropy*, if

$$\Pr_{(\mathsf{pk},\mathsf{sk}) \leftarrow \mathsf{IGen}(\lambda)} (H_\infty(W|(W, St) \leftarrow P_1(\mathsf{sk})) \geq \alpha) \geq 1 - 2^{-\alpha}.$$

Equivalently, except with probability $2^{-\alpha}$ over the choice of keys, the min-entropy of W will be at least α.

Definition 6 (Lossy identification). *A canonical identification scheme is lossy if there exists a lossy key generation algorithm* $\mathsf{LossyIGen}$ *that takes the security parameter* λ *as input and returns a public key* $\mathsf{pk}_{\mathsf{ls}}$ *without any secret key. A lossy* ID *scheme satisfied two security properties:*

1. *Indistinguishability of keys: the public keys generated by* IGen *and* $\mathsf{LossyIGen}$ *are indistinguishable. Equivalently, for any quantum adversary* A*, the following distinguishing advantage is negligible*

$$\begin{aligned} Adv^{\mathsf{loss}}_{\mathsf{ID}}(A) :=& |\Pr(A(\mathsf{pk}_{\mathsf{ls}}) = 1 | \mathsf{pk}_{\mathsf{ls}} \leftarrow \mathsf{LossyIGen}(\lambda)) \\ & - \Pr(A(\mathsf{pk}) = 1 | (\mathsf{pk}, \mathsf{sk}) \leftarrow \mathsf{IGen}(\lambda))|. \end{aligned}$$

2. *Statistical lossy soundness: given a lossy public key* pk*, not even an unbounded quantum adversary can impersonate the prover. A lossy* ID *scheme has* ϵ_{ls}*-lossy soundness if for any quantum adversary the probability to impersonate the prover is bounded by* ϵ_{ls}*.*

Definition 7 (No-abort honest-verifier zero-knowledge). *A canonical identification scheme* ID *is* $\varepsilon_{\mathsf{zk}}$*-perfect no-abort honest-verifier zero-knowledge (*$\varepsilon_{\mathsf{zk}}$*-perfect* $\mathsf{na\text{-}HVZK}$*) if there exists a* ppt *algorithm* Sim *which given only the public key* pk*, outputs* (W, c, Z) *such that:*

- *The statistical distance between* $(W, c, Z) \leftarrow \mathsf{Sim}(\mathsf{pk})$ *and* $(W, c, Z) \leftarrow \mathsf{Trans}(\mathsf{pk})$ *is at most* $\varepsilon_{\mathsf{zk}}$*;*

- *The distribution of* c *in* $(W, c, Z) \leftarrow \mathsf{Sim}(\mathsf{pk})$ *conditioned on* $c \neq \perp$ *is uniform in* ChSet.

Finally, an identification scheme has *computational unique response* (CUR) property if it is computationally infeasible to produce two different valid responses Z, Z' for any commitment/challenge pair (W, c). In the context of lossy identification scheme, the computational unique response property is required to ensure the strong unforgeability of the signature scheme.

D Signatures

Definition 8 (Digital signature). *A digital signature scheme* SIG *consists of a triple of* ppt *algorithms* $(\mathsf{Gen}, \mathsf{Sign}, \mathsf{Ver})$.

- *The key generation algorithm* $\mathsf{Gen}(\lambda)$ *inputs the security parameter* λ *and returns the public and secret keys* $(\mathsf{vk}, \mathsf{sk})$. *The* vk *defines the message space for messages* μ.
- *The signing algorithm* $\mathsf{Sign}(\mathsf{sk}, \mu)$ *returns a signature* σ.
- *The deterministic verification algorithm* $\mathsf{Ver}(\mathsf{vk}, \mu, \sigma)$ *returns* 1 *for accept or* 0 *for reject.*

A signature scheme admits correctness error $\delta \geq 0$, if for every pair of outputs $(\mathsf{sk}, \mathsf{vk})$ of $\mathsf{Gen}(\lambda)$ and any message μ in the message space, we have

$$\Pr[\mathsf{Ver}(\mathsf{vk}, \mu, \mathsf{Sign}(\mathsf{sk}, \mu)) = 0] \leq \delta,$$

where the probability is taken over the randomness of algorithms Sign.

Definition 9 (Unforgeability). *A signature scheme* $\mathsf{SIG} := (\mathsf{Gen}, \mathsf{Sign}, \mathsf{Ver})$ *is said to be unforgeable against one-per-message chosen message attack (*$\mathsf{UF\text{-}CMA}_1$*) in the quantum random oracle model if: for every* ppt *quantum forger* $\mathcal{F}$ *having quantum access to the random oracle and classical access to the signing oracle, the forging probability that, after seeing the public key and* $Q = \mathrm{poly}(n)$ *adaptively chosen distinct messages* M_i *of his choice and their signatures* $\{(M_i, \mathsf{Sign}(\mathsf{sk}, M_i))\}_{i=1,\cdots,Q}$, *the forger* $\mathcal{F}$ *can produce* $M^* \notin \{M_i\}_{i=1,\cdots,Q}$ *and* σ^* *such that* $\mathsf{Ver}(\mathsf{vk}, M^*, \sigma^*) = 1$, *is negligibly small. The forging probability is taken over the randomness of* $\mathsf{Gen}, \mathsf{Sign}$ *and* $\mathcal{F}$, *and denoted as* $Adv^{\mathsf{UF\text{-}CMA}_1}_{\mathsf{SIG}}(\mathcal{F})$.

One can extend the definition to the case where the forger may obtain more than one signature for any of Q adaptively chosen messages $\{M_i\}$. If no quantum forger $\mathcal{F}$ can produce a valid signature for a message $M^* \notin \{M_i\}_{i=1,\cdots,Q}$, we say the signature scheme is unforgeable against chosen message attack ($\mathsf{UF\text{-}CMA}$).

In the *strong* $\mathsf{UF\text{-}CMA}/\mathsf{UF\text{-}CMA}_1$ setup (denoted $s\mathsf{UF\text{-}CMA}/s\mathsf{UF\text{-}CMA}_1$), the adversary may return a forgery for a message which has already been queried previously, but with a different signature.

E Fiat-Shamir transform

One can construct a signature scheme from an identification scheme where the hardness of the signature inherits from the ID scheme. Let $\mathsf{ID} := (\mathsf{IGen}, \mathsf{P}, \mathsf{V})$ be a canonical identification scheme. One obtains the signature scheme $\mathsf{SIG} := (\mathsf{G} = \mathsf{IGen}, \mathsf{Sign}, \mathsf{Ver})$ via the Fiat-Shamir transformation. Such approach uses the Fiat-Shamir transform as illustrated in Figure 6, which can be used to construct lattice-based signatures without trapdoor [32,33].

Algorithm $\mathsf{Sign}(\mathsf{sk}, \mu)$	Algorithm $\mathsf{Ver}(\mathsf{pk}, \sigma, \mu)$
$i = 0$	Parse $\sigma := (W, Z)$
while $Z = \perp$ **and** $i \leq \tau_m$ **do**	$c = H(W \| \mu)$
$\quad i := i + 1$	**return** $\mathsf{V}(\mathsf{pk}, W, c, Z) \in \{0, 1\}$
$\quad (W, St) \leftarrow \mathsf{P}_1(\mathsf{sk})$	
$\quad c = H(W \| \mu)$	
$\quad Z \leftarrow \mathsf{P}_2(\mathsf{sk}, W, c, St)$	
if $Z = \perp$ return $\sigma = \perp$	
output $\sigma := (W, Z)$	

Fig. 6. Fiat-Shamir signature obtained from an $\mathsf{ID} = (\mathsf{IGen}, \mathsf{P}, \mathsf{V})$

Such transform was known to be secure in the random oracle model in the classic setting, which does not automatically imply security in the quantum setup. Kiltz, Lyubashevsky and Schaffner [30] presents a generic framework for constructing tight reductions in the quantum random-oracle (QROM) model, which can be constructed from a lossy ID scheme. In particular, if the underlying identification scheme is lossy and $\mathsf{na\text{-}HVZK}$, then the $\mathsf{UF\text{-}CMA}_1$ security of the scheme is tightly based on the hardness of distinguishing regular and lossy public keys of the identification scheme. We will use Theorem 3.1 of [30] in this work and cite it as follows.

Theorem 3.1 of [30]. *Consider an identification scheme* ID *which is lossy,* $\varepsilon_{\mathsf{zk}}$*-perfect* $\mathsf{na\text{-}HVZK}$*, has* α *bits of entropy and is* $\varepsilon_{\mathsf{ls}}$*-lossy sound and the signature scheme* SIG *obtained by applying the Fiat-Shamir transform to the identification scheme* ID*, as in Figure 6. For any quantum adversary* A *against* $\mathsf{UF\text{-}CMA}_1$ *(and* $s\mathsf{UF\text{-}CMA}_1$*) security that issues at most* Q_H *quantum queries to the random oracle and* Q_S *classical signing queries, there exists a quantum adversary* B *against* ID *(and a quantum adversary* C *against* CUR*) such that*

$$Adv^{\mathsf{UF\text{-}CMA}_1}_{\mathsf{SIG}}(A) \leq Adv^{\mathsf{loss}}_{\mathsf{ID}}(B) + 8(Q_H + 1)^2 \cdot \varepsilon_{\mathsf{ls}} + \tau_m \cdot Q_S \cdot \varepsilon_{\mathsf{zk}} + 2^{-\alpha+1}$$
$$Adv^{\mathsf{SUF\text{-}CMA}_1}_{\mathsf{SIG}}(A) \leq Adv^{\mathsf{loss}}_{\mathsf{ID}}(B) + 8(Q_H + 1)^2 \cdot \varepsilon_{\mathsf{ls}} + \tau_m \cdot Q_S \cdot \varepsilon_{\mathsf{zk}} + 2^{-\alpha+1} + Adv^{\mathsf{CUR}}_{\mathsf{ID}}(C)$$

and $Time(B) = Time(C) = Time(A) + \tau_m \cdot Q_H \approx Time(A)$.

It can be noted that $\mathsf{UF\text{-}CMA}_1$-secure signatures can be de-randomized with a pseudo-random function PRF to obtain a $\mathsf{UF\text{-}CMA}$ secure signatures with deterministic signing [11] and such reduction is tight. Thus one can de-randomize the signature scheme in Figure 6 by using a PRF to obtain a deterministic signature DSIG. Then for any quantum adversary A against the $\mathsf{UF\text{-}CMA}$ security of DSIG that issues at most Q_H quantum queries to the random oracle and Q_S classical signing queries, there exists a quantum adversary B against ID and a quantum adversary D against the PRF such that

$$\mathrm{Adv}^{\mathsf{UF\text{-}CMA}}_{\mathsf{DSIG}}(A) \leq \mathrm{Adv}^{\mathsf{loss}}_{\mathsf{ID}}(B) + 8(Q_H+1)^2 \cdot \varepsilon_{\mathsf{ls}} + \tau_m \cdot Q_S \cdot \varepsilon_{\mathsf{zk}} + 2^{-\alpha+1} + \mathrm{Adv}^{\mathsf{PR}}_{\mathsf{PRF}}(D)$$
$$\mathrm{Adv}^{\mathsf{SUF\text{-}CMA}}_{\mathsf{DSIG}}(A) \leq \mathrm{Adv}^{\mathsf{loss}}_{\mathsf{ID}}(B) + 8(Q_H+1)^2 \cdot \varepsilon_{\mathsf{ls}} + \tau_m \cdot Q_S \cdot \varepsilon_{\mathsf{zk}} + 2^{-\alpha+1} + \mathrm{Adv}^{\mathsf{PR}}_{\mathsf{PRF}}(D) + \mathrm{Adv}^{\mathsf{CUR}}_{\mathsf{ID}}(C)$$

where the $\mathrm{Adv}^{\mathsf{PR}}_{\mathsf{PRF}}(D)$ denotes the distinguishing advantage of the adversary D, w.r.t a perfect random function PR.

F Signature scheme from the lossy identification and security proofs

The signature is constructed from the identification scheme in Figure 1 by using the Fiat-Shamir transform. Theorem 3.1 on page 25 states the upper bound for the security of our signature. We instatiate it with concrete parameters in Section 5.

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

1 : Sample $\mathbf{f} = \{f_i\}_{i=1}^{d}$ and t, where $f_i \hookleftarrow U_\beta^n$ and $t \hookleftarrow U_{R_q}$

2 : Compute $\mathbf{h} = (h_1, \ldots, h_{d-1}, 1)$ such that $\sum_{i=1}^{d} h_i f_i \equiv t \pmod q$

3 : **return** $\mathsf{pk} := (\mathbf{h}, t)$ and $\mathsf{sk} := \mathbf{f}$

Algorithm $\mathsf{Sign}(\mathsf{sk}, \mu)$:

4 : Sample $\mathbf{y} = \{y_i\}_{i=1}^{d-1}$ where $y_i \hookleftarrow U_\gamma^n$

5 : Compute $c = H\left(\left\lfloor \sum_{i=1}^{d-1} h_i y_i \pmod q \right\rceil_r, \mu\right)$

6 : Compute $\mathbf{z} = \{z_i\}_{i=1}^{d-1}$ where $z_i = y_i + c \cdot f_i$

7 : Compute $w = \sum_{i=1}^{d-1} h_i z_i - c \cdot f_d$

8 : **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**

9 : **return** $\perp$

10 : **return** $(\mathbf{z}, c)$

Algorithm $\mathsf{Ver}(\mathsf{pk}, \mu, \mathbf{z}, c)$:

11 : **if** $\forall\, 1 \leq i \leq d-1, \|z_i\|_\infty < \gamma - \beta \cdot \kappa$ **and**
$H\left(\left\lfloor \sum_{i=1}^{d-1} h_i z_i - t \cdot c \pmod q \right\rceil_r, \mu\right) = c$ **then**

12 : **return** Accept

13 : **return** Reject

Fig. 7. A signature scheme obtained from the lossy identification scheme.

Proofs for signature scheme in Section 3 In this subsection, we show the ID scheme admits properties such as $\mathsf{na\text{-}HVZK}$, correctness, lossy, min-entropy and CUR. For convenience, we denote $B := \beta \cdot \kappa$.

Proof of Lemma 1 : We sketch the proof. Our choice of parameters guarantees that the infinity norm of z_i in the output is bounded by $\gamma - B$. It is clear that each z_i bounded by $\gamma - B$ has an equal probability of being generated. Furthermore, the probability that some $\mathbf{z}$ such that $\|\mathbf{z}\|_\infty \leq \gamma - B$ is generated is precisely $(|U_{\gamma-B}|/|U_\gamma|)^{n(d-1)}$, as in Line 9 of Figure 2. Finally, we note that Line 12 of Figure 2 satisfies $\sum_{i=1}^{d-1} h_i z_i - t \cdot c = \sum_{i=1}^{d-1} h_i y_i - c \cdot f_d \pmod q$. Thus, the step in Line 7 of Algorithm Trans is identical to that of Line 13 of Algorithm Sim. □

Proof of Lemma 2 : It can be checked that when the output is not $\perp$, the verification procedure will always accept. This is due to the conditions in Line 9 and 12 of Figure 1.

Furthermore, Lemma 1 shows Algorithms Trans and Sim output $\perp$ with the same probability. Thus, it suffices to focus on Algorithm Sim.

First, the probability of producing some $\mathbf{z}$, that is to say Line 9 of Algorithm Sim does not return $\perp$, is

$$\left(\frac{2(\gamma - B) + 1}{2\gamma + 1}\right)^{n(d-1)} \approx \left(1 - \frac{B}{\gamma}\right)^{n(d-1)} \approx \exp\left(-Bn(d-1)/\gamma\right).$$

Second, we heuristically assume the low-bits $[w']_r$ are uniform, where $w' = \sum_{i=1}^{d-1} h_i z_i - t \cdot c \pmod q$ for a uniform $\mathbf{z}$.

Thus, the probability that $\|[w']_r\|_\infty < 2^{r-1} - B$ is about $\left(-Bn/2^{r-1}\right)$.

Finally, we will also assume that w' is uniform mod q and thus the probability that $\|w'\|_\infty < \lfloor q/2 \rfloor - B$ is about $\exp\left(-2Bn/q\right)$.

In these estimated inequalities, we assumed $q \gg 2^{r-1} \gg \gamma \gg B = \beta \cdot \kappa$. The overall acceptance probability is as stated. □

Proof of Lemma 3 : We will sketch the proof, which follows the same framework of [30].

We now introduce the *impersonation game*. Consider an unbounded quantum adversary $\mathcal{A}$ who receives the lossy key $(\mathbf{h}, t)$ produced from $\mathsf{LossyIGen}$. The adversary $\mathcal{A}$ attempts to impersonate with the following steps: outputting u, taking c uniformly from $\mathcal{B}_\kappa$, and then producing $\mathbf{z}$. We will prove that for almost all lossy keys generated from $\mathsf{LossyIGen}$, for any u, there exists at most one possible c that allows the adversary $\mathcal{A}$ to win. Since c is taken uniformly from $\mathcal{B}_\kappa$, this implies the adversary wins at a chance of at most $\frac{1}{|\mathcal{B}_\kappa|}$, for almost all lossy keys.

Now, suppose that for some u, there exist pairs $(\mathbf{z}, c)$ and $(\mathbf{z}', c')$, such that $c \neq c'$, in which the adversary $\mathcal{A}$ wins. Thus,

$$u = \left\lfloor \sum_{i=1}^{d-1} h_i z_i - t \cdot c \right\rceil_r = \left\lfloor \sum_{i=1}^{d-1} h_i z_i' - t \cdot c' \right\rceil_r.$$

Using the bounds on z_i's and rounding properties, there exists some w such that

$$\sum_{i=1}^{d-1} h_i (z_i - z_i') + w = t \cdot (c - c'). \tag{2}$$

We have the bounds $\|w\|_\infty \leq 2^r$, $\|z_i - z_i'\|_\infty \leq 2(\gamma - B)$ and $\|c - c'\|_\infty \leq 2$. To prove this only happens rarely, we will show the following statement that, over random choices of $(\mathbf{h}, t)$, the equation $\sum_{i=1}^{d-1} h_i z_i^* + w^* = t \cdot c^*$ is satisfied (i.e. there exist such $\mathbf{z}^* = \{z_i^*\}_i, w^*$ and c^*) with probability bounded by

$$2 \cdot |C| \cdot \frac{(4(\gamma - B) + 1)^{n(d-1)} \cdot (2^{r+1} + 1)^n}{q^n}. \tag{3}$$

Here, for convenience, we denote sets $Z = \{\forall i \in [1, d-1],\ z_i^* \in R_q \mid \|z_i^*\|_\infty \leq 2(\gamma - B)\}$, $W = \{w^* \in R_q \mid \|w^*\|_\infty \leq 2^r\}$ and $C = \{c^* \in R_q \mid \|c^*\|_\infty \leq 2\}$. Note that one can also restrict C to contain elements c^* of at most 2κ non-zero entries.

We consider two cases depending on whether $\mathbf{z}^* = \mathbf{0}$ or not. First, suppose $\mathbf{z}^* = \mathbf{0}$. We consider

$$\Pr_{t \hookleftarrow R_q} \left[\exists\, (w^*, c^*) \in R_q \times C \text{ s.t. } w^* = t \cdot c^*\right]$$

By [36, Lemma 2.2], the elements c^* of norm $< \sqrt{q/2}$ are invertible in R_q as we choose $q \equiv 5 \pmod 8$. Thus, the above probability is

$$\begin{aligned}\Pr_{t \hookleftarrow R_q} \left[\exists\, (w^*, c^*) \in R_q \times C \text{ s.t. } w^* \cdot (c^*)^{-1} = t\right] &\leq \sum_{w^* \in W, c^* \in C} \Pr_{t \hookleftarrow R_q} \left[t = w^* \cdot (c^*)^{-1}\right] \\ &\leq |C| \cdot \frac{(2^{r+1}+1)^n}{q^n}\end{aligned}$$

Second, suppose that $\mathbf{z}^* \neq \mathbf{0}$. For some fixed $(\mathbf{z}^*, w^*, c^*)$, we assume that $z_1 \neq 0$; otherwise, take any non-zero z_i. We have

$$\begin{aligned}\Pr_{\forall i,\, h_i \hookleftarrow R_q, t \hookleftarrow R_q} \left[\sum_{i=1}^{d-1} h_i z_i^* + w^* = t \cdot c^*\right] &= \Pr_{h_1 \hookleftarrow R_q} \left[h_1 = \frac{t \cdot c^* - w^* - \sum_{i=2}^{d-1} h_i z_i^*}{z_1}\right] \\ &\leq \frac{1}{q^n}.\end{aligned}$$

The invertibility of z_1 is due to the small norm of z_1 by [36, Lemma 2.2] again. By the union bound, the probability over all $(\mathbf{z}^* \neq \mathbf{0}, w^*, c^*)$ can be bounded by

$$\sum_{\mathbf{z}^* \in Z \setminus \{\mathbf{0}\}, w^* \in W, c^* \in C} \frac{1}{q^n} \leq |C| \cdot \frac{(4(\gamma - B) + 1)^{n(d-1)} \cdot (2^{r+1}+1)^n}{q^n}.$$

Thus, Equation (2) is satisfied when $(\mathbf{h}, t)$ is sampled from $\mathsf{LossyIGen}$, with probability bounded by the one given in Equation (3).

Finally, as c is taken uniformly from $\mathcal{B}_\kappa$ in the impersonation game, the adversary wins at a chance of at most

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

We note that $|C| \leq |\mathcal{B}_\kappa|^2$. □

Proof of Lemma 4 : We will bound the probability that the commitment u is unique, averaged over a random choice $\mathbf{h}'$ - e.g.,

$$\Pr_{\mathbf{h}' \hookleftarrow U_{R_q}^{d-1}} \left[\exists\, \mathbf{y} \neq \mathbf{y}', \left\lfloor \sum_{i=1}^{d-1} h_i y_i \right\rceil_r = \left\lfloor \sum_{i=1}^{d-1} h_i y_i' \right\rceil_r\right].$$

If there exist distinct $\mathbf{y}$ and $\mathbf{y}'$ such that $\left\lfloor \sum_{i=1}^{d-1} h_i y_i \right\rceil_r = \left\lfloor \sum_{i=1}^{d-1} h_i y_i' \right\rceil_r$, then there exist non-zero elements $\mathbf{y}^*$ and w^* such that $\sum_{i=1}^{d-1} h_i y_i^* = w^* \pmod q$,

where $\|\mathbf{y}^*\|_\infty \le 2\gamma$ and $\|w^*\|_\infty \le 2^r$. Since $\mathbf{y}^*$ is nonzero, we may assume that $y_1^* \neq 0$ w.l.o.g. Again, the elements y_1^* of small norm are invertible by [36, Lemma 2.2]. Thus, we have

$$\sum_{\mathbf{y}^*, w^*} \Pr_{\mathbf{h}' \hookleftarrow U_{R_q}^{d-1}} \left[\sum_{i=1}^{d-1} h_i y_i^* = w^* \right] = \sum_{\mathbf{y}^*, w^*} \Pr_{h_1 \hookleftarrow U_{R_q}} \left[h_1 = \frac{w^* - \sum_{i=2}^{d-1} h_i y_i^*}{y_1^*} \right] \le \frac{(4\gamma+1)^{n(d-1)}(2^{r+1}+1)^n}{q^n}$$

Denote $\delta := (4\gamma+1)^{n(d-1)}(2^{r+1}+1)^n q^{-n}$. With probability at least $1-\delta$ over the choices of public key $\mathbf{h}'$, the min entropy of $\left\lfloor \sum_{i=1}^{d-1} h_i y_i \pmod q \right\rceil_r$ is at least $(d-1)n\log(2\gamma+1)$. Therefore, the identification has α bits of min-entropy, as stated in the lemma. □

Proof of Lemma 5 : Suppose an adversary can produce two transcripts $(u, c, \mathbf{z})$, $(u, c, \mathbf{z}')$ such that $\mathbf{z} \neq \mathbf{z}'$ and $\mathsf{V}(\mathsf{pk}, u, c, \mathbf{z}) = \mathsf{V}(\mathsf{pk}, u, c, \mathbf{z}') = 1$. We see that there exist $\mathbf{z}^* \neq \mathbf{0}$ and w^* such that $\sum_{i=1}^{d-1} h_i z_i^* = w^*$, where $\|\mathbf{z}^*\|_\infty \le 2(\gamma - B)$ and $\|w^*\|_\infty \le 2^r$. One can bound the probability, over the choice of $\mathbf{h}$, that there exists such $(\mathbf{z}^*, w^*)$ using a similar method as in Lemma 3 or Lemma 4. □

G Proofs for BLISS-like signature

Proof of Lemma 6 : The only difference between the genuine signature algorithm and Hybrid 1 is the necessary programming of the random oracle at $t := \langle \mathbf{a}, \mathbf{z} \rangle + q\,c = \langle \mathbf{a}, \mathbf{y} \rangle \pmod{2q}$, without checking whether the value t was previously set or queried. We bound such probability for any t and $\mathbf{a}$. We have

$$\Pr_{\forall i,\, y_i \hookleftarrow \mathcal{D}_\sigma^n}[\langle \mathbf{a}, \mathbf{y} \rangle = t] = \Pr_{\forall i,\, y_i \hookleftarrow \mathcal{D}_\sigma^n} \left[\sum_{i=1}^{d-1} 2h_i y_i + (q-2)\, y_d = t \right] \le \max_{w \in R_q} \Pr_{y_d \hookleftarrow \mathcal{D}_\sigma^n}[y_d = w] \le 2^{-n}.$$

The total number of values programmed will be bounded by $q_H + q_S$. Thus, the probability that we have any collision with one of the previous calls of $\mathcal{O}_{\mathsf{Sign}}$ is bounded by $q_S(q_H + q_S)2^{-n}$. This provides an upper bound on the advantage of the distinguisher algorithm $\mathcal{D}$. □

Proof of Theorem 1 : Let $\mathbf{h}' = \{h_i\}_{i=1}^{d-1}$ be the instance sampled from the the variant $\mathsf{MNTRU}_{q,n,d,D_1,D_2}$ distribution, where D_1 and D_2 sample polynomials from $D_{\mathbb{Z},\sigma}^n$. We construct the public key $\mathbf{a}$ for the signature in Hybrid 2 such that $\mathbf{a} = (2h_1, \cdots, 2h_{d-1}, q-2)$ in R_{2q}^d. We aim to find a vector $\mathbf{x}$ such that $\langle \mathbf{a}, \mathbf{x} \rangle = 0 \pmod q$, which implies the solution $\langle \mathbf{h}, \mathbf{x} \rangle = 0 \pmod q$.

When the adversary $\mathcal{A}$ wants to see the signature on some message, one calls the signing algorithm from Hybrid 2. When the random oracle is queried during

signing or random oracle access, one programs the random oracle, maintaining a list of all queries to keep track of the same query being made. Finally, $\mathcal{A}$ produces a forgery $(\mathbf{z}, c)$ on a message μ, using at most q_S signing queries to Hybrid 2 and q_H hash queries to the random oracle H, and succeeds with probability δ. Write $q_T = q_S + q_H$ as the upper bound on the number of times the random oracle is called or programmed during $\mathcal{A}$'s attack.

We denote $\beta := \binom{n}{\kappa}$. With probability $1 - 1/\beta$, the forgery relies on having had c as the result of a query (hash or sign) previously. With probability $\delta - 1/\beta$, this leads to two possible sources of the forgery: collision from the Hybrid 2 and collision from the random oracle query. We cover these in two cases. First, suppose c was obtained from a signing oracle query. Then, there exist pairs $(\mathbf{z}, \mu)$ and $(\mathbf{z}', \mu')$ such that $H(\langle \mathbf{a}, \mathbf{z} \rangle + q \cdot c, \mu) = H(\langle \mathbf{a}, \mathbf{z}' \rangle + q \cdot c, \mu')$, which implies

$$\mu = \mu' \text{ and } \langle \mathbf{a}, \mathbf{z} \rangle + q \cdot c = \langle \mathbf{a}, \mathbf{z}' \rangle + q \cdot c \pmod{2q},$$

due to the collision resistance. This yields $\sum_{i=1}^{d-1} h_i(z_i - z_i') - (z_d - z_d') \equiv 0 \mod q$. Note that $\|\mathbf{z} - \mathbf{z}'\|$ is bounded by $2\eta\sigma\sqrt{nd}$ and $\mathbf{z} \neq \mathbf{z}'$. Furthermore, $\mathbf{z} \not\equiv \mathbf{z}' \pmod{q}$, since $\|\mathbf{z}\|_\infty$ and $\|\mathbf{z}'\|_\infty \leq q/4$. Thus, $\mathbf{z} - \mathbf{z}'$ is a non-zero solution of the variant $\mathsf{MNTRU}_{q,n,d,D_1,D_2,B}$.

Second, suppose c was obtained from a hash oracle query, the Forking Lemma in [10,18] states that, with probability $(\delta - \frac{1}{\beta})(\frac{\delta - \beta^{-1}}{q_T} - \frac{1}{\beta}) \approx \delta^2/q_T$, there exist $(\mathbf{z}, c)$ and $(\mathbf{z}', c')$ with $c \neq c'$ such that

$$\sum_{i=1}^{d-1} 2h_i(z_i - z_i') - (q-2)(z_d - z_d') \equiv q(c - c') \mod 2q.$$

This yields $\sum_{i=1}^{d-1} h_i(z_i - z_i') - (z_d - z_d') \equiv 0 \mod q$. The c, c' are binary and hence $c \neq c' \pmod 2$. Given that $\|\mathbf{z}\|_\infty$ and $\|\mathbf{z}'\|_\infty \leq q/4$, we see that $\mathbf{z} \not\equiv \mathbf{z}' \pmod q$ and hence $\mathbf{z} - \mathbf{z}'$ is a nonzero solution of the variant $\mathsf{MNTRU}_{q,n,d,D_1,D_2,B}$. □

H Lattice reduction estimates

In MNTRU lattices, the secret $(\mathbf{f}, g)$ has norm less than the Gaussian heuristic norm of the shortest non-zero vector in $\Lambda_{\mathbf{h}}$ and leads to the so-called unique-SVP problems (uSVP).

We recall the estimate for solving uSVP by Gama and Nguyen [22]. Generally, the work in [22] showed that the shortest vector in the uSVP_γ problem can be recovered as soon as $\gamma \geq \tau \cdot \delta^m$, where δ is the root Hermite factor of the reduction algorithm, and γ is the ratio gap of λ_2/λ_1 in uSVP lattices. Here $\tau < 1$ is an empirical constant determined by experiments: it has been investigated that τ lies in between 0.3 and 0.4 when using the BKZ algorithm [4]. Ordinarily, the second minimum λ_2 of a uSVP lattice is approximated by using the Gaussian heuristic to predict the norm of the shortest vector in random lattices. The δ is a decreasing function of β and therefore we want to maximize δ.

In the New Hope key exchange paper [7], another alternative method for estimating the cost for solving uSVP is given, and it has been investigated extensively by Albrecht et al. in work [4]. Instead of looking at the gap of the uSVP directly, it considers the evolution of the Gram-Schmidt coefficients of the unique shortest vector in the BKZ tours. More precisely, it compares the expected length of the projected (expected) shortest vector $\mathbf{v}$ with the Gram-Schmidt lengths estimated by the Geometric Series Assumption (GSA) [48].

Geometric Series Assumption (GSA) Let k be a rank of a lattice Λ. For $1 \leq i \leq k-1$, the norm of the Gram-Schmidt vectors $\mathbf{b}_i^*$ in the lattice reduction basis BKZ_β satisfy

$$\|\mathbf{b}_i^*\| = \delta_\beta^2 \cdot \left\|\mathbf{b}_{i+1}^*\right\|.$$

The main idea is that partial information of the shortest vector $\mathbf{v}$ will be recovered in the last block, when the orthogonal projection of $\mathbf{v}$ to the first $d-\beta$ Gram-Schmidt vectors is shorter than the expected $\mathbf{b}_{d-\beta+1}^*$ predicted by the GSA assumption. Thus the success condition for recovering the secret for MNTRU problems, can be formulated as follows:

$$\sqrt{\beta/\dim(\mathcal{L})} \cdot \left\|\left[\mathbf{f}, g\right]\right\| \leq \delta_\beta^{2\beta-\dim(\mathcal{L})-1} \cdot \mathrm{Vol}(\mathcal{L})^{1/\dim(\mathcal{L})}.$$