Learning a Zonotope and More: Cryptanalysis of NTRUSign Countermeasures

Léo Ducas, Phong Q. Nguyen

Learning a Zonotope and More: Cryptanalysis of NTRUSign Countermeasures

Léo Ducas and Phong Q. Nguyen

[1] ENS, Dept. Informatique, 45 rue d'Ulm, 75005 Paris, France.
http://www.di.ens.fr/~ducas/
[2] INRIA, France and Tsinghua University, Institute for Advanced Study, China.
http://www.di.ens.fr/~pnguyen/

Abstract NTRUSign is the most practical lattice signature scheme. Its basic version was broken by Nguyen and Regev in 2006: one can efficiently recover the secret key from about 400 signatures. However, countermeasures have been proposed to repair the scheme, such as the perturbation used in NTRUSign standardization proposals, and the deformation proposed by Hu *et al.* at IEEE Trans. Inform. Theory in 2008. These two countermeasures were claimed to prevent the NR attack. Surprisingly, we show that these two claims are incorrect by revisiting the NR gradient-descent attack: the attack is more powerful than previously expected, and actually breaks both countermeasures in practice, *e.g.* 8,000 signatures suffice to break NTRUSign-251 with one perturbation as submitted to IEEE P1363 in 2003. More precisely, we explain why the Nguyen-Regev algorithm for learning a parallelepiped is heuristically able to learn more complex objects, such as zonotopes and deformed parallelepipeds.

1 Introduction

There is growing interest in cryptography based on hard lattice problems (see the survey [22]). The field started with the seminal work of Ajtai [2] back in 1996, and recently got a second wind with Gentry's breakthrough work [7] on fully-homomorphic encryption. It offers asymptotical efficiency, potential resistance to quantum computers and new functionalities. There has been significant progress in provably-secure lattice cryptography in the past few years, but from a practical point of view, very few lattice schemes can compete with standardized schemes for now. This is especially true in the case of signature schemes, for which there is arguably only one realistic lattice alternative: NTRUSign [11], which is an optimized instantiation of the Goldreich-Goldwasser-Halevi (GGH) signature scheme [9] using the compact lattices introduced in NTRU encryption [14] and whose performances are comparable with ECDSA. By comparison, signatures have size beyond 10,000 bits (at 80-bit security level) for the most efficient provably-secure lattice signature scheme known, namely the recent scheme of Lyubashevsky [19].

However, NTRUSign has no provable-security guarantee. In fact, the GGH signature scheme and its simplest NTRUSign instantiation were broken at EUROCRYPT '06 by Nguyen and Regev [23], who presented a polynomial-time

key-recovery attack using a polynomial number of signatures: in the case of NTRUSign, 400 signatures suffice in practice to disclose the secret key within a few hours. In the GGH design, a signature is a lattice point which is relatively close to the (hashed) message. Clearly, many lattice points could be valid signatures, but GGH selects one which is closely related to the secret key: each message–signature pair actually discloses a sample almost uniformly distributed in a secret high-dimensional parallelepiped. The NR attack works by learning such a parallelepiped: given a polynomial number of samples of the form $\sum_{i=1}^{n} x_i \mathbf{b}_i$ where the x_i's are picked uniformly at random from $[-1, 1]$ and the secret vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{R}^n$ are linearly independent, the attack recovers the parallelepiped basis $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$, by finding minima of a certain multivariate function, thanks to a well-chosen gradient descent. The NR attack motivated the search of countermeasures to repair NTRUSign:

- The very first countermeasure already appeared in half of the parameter choices of NTRU's IEEE P1363.1 standardization proposal [17], the other half being broken by NR. It consists of applying the signature generation process twice, using two different NTRU lattices, the first one being kept secret: here, the secret parallelepiped becomes the Minkowski sum of two secret parallelepipeds, which is a special case of zonotopes. This slows down signature generation, and forces to increase parameters because the signature obtained is less close to the message. However, no provable security guarantee was known or even expected. In fact, heuristic attacks have been claimed by both the designers of NTRUSign [10] and more recently by Malkin *et al.* [20], but both are impractical: the most optimistic estimates [10,20] state that they both require at least 2^{60} signatures, and none have been fully implemented. Yet, as a safety precaution, the designers of NTRUSign [10] only claim the security of NTRUSign with perturbation up to 1 million signatures in [11]. Still, breaking this countermeasure was left as an open problem in [23].
- In 2008, Hu, Wang and He [16] proposed a simpler and faster countermeasure in IEEE Trans. Inform. Theory, which we call IEEE-IT, where the secret parallelepiped is deformed. Again, the actual security was unknown.
- Gentry, Peikert and Vaikuntanathan [8] proposed the first provably secure countermeasure for GGH signatures, by using a randomized variant [18] of Babai's nearest plane algorithm. However, this slows down signature generation significantly, and forces to increase parameters because the signatures obtained are much less close to the message. As a result, the resulting signature for NTRUSign does not seem competitive with classical signatures: no concrete parameter choice has been proposed.

Our Results. We revisit the Nguyen-Regev gradient-descent attack to show that it is much more powerful than previously expected: in particular, an optimized NR attack can surprisingly break in practice both NTRU's perturbation technique [11] as recommended in standardization proposals [17,13], and the IEEE-IT countermeasure [16]. For instance, we can recover the NTRUSign secret key in a few hours, using 8,000 signatures for the original NTRUSign-251

scheme with one perturbation submitted to IEEE P1363 standardization in 2003, or only 5,000 signatures for the latest 80-bit-security parameter set [13] proposed in 2010. These are the first successful experiments fully breaking NTRUSign with countermeasures. Note that in the perturbation case, we have to slightly modify the original NR attack. The warning is clear: our work strongly suggests to dismiss all GGH/NTRUSign countermeasures which are not supported by some provable security guarantee.

Our work sheds new light on the NR attack. The original analysis of Nguyen and Regev does not apply to any of the two NTRUSign countermeasures, and it seemed *a priori* that the NR attack would not work in these cases. We show that the NR attack is much more robust than anticipated, by extending the original analysis of the Nguyen-Regev algorithm for learning a parallelepiped, to tackle more general objects such as zonotopes (to break the NTRUSign countermeasure with a constant number of perturbations) or deformed parallelepipeds (to break the IEEE-IT countermeasure). For instance, in the zonotope case, the parallelepiped distribution $\sum_{i=1}^{n} x_i \mathbf{b}_i$ is replaced by $\sum_{i=1}^{m} x_i \mathbf{v}_i$ where $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ are secret vectors with $m \geq n$. The key point of the NR attack is that all the local minima of a certain multivariate function are connected to the directions $\mathbf{b}_i$'s of the secret parallelepiped. We show that there is somewhat a similar (albeit more complex) phenomenon when the parallelepiped is replaced by zonotopes or deformed parallelepipeds: there, we establish the existence of local minima connected to the secret vectors spanning the object, but we cannot rule out the existence of other minima. Yet, the attack works very well in practice, as if there were no other minima.

Roadmap. In Sect. 2, we recall background on NTRUSign and the NR attack. In Sect. 3, we attack NTRU's perturbation countermeasure, by learning a zonotope. In Sect. 4, we attack the IEEE-IT countermeasure, by learning a deformed parallelepiped. More information is provided in the full version [5].

2 Background and Notation

2.1 Notation

Sets. $\mathbb{Z}_q$ is the ring of integers modulo q. $\mathbb{N}$ and $\mathbb{Z}$ denote the usual sets. $[n]$ denotes $\{1, \cdots, n\}$. $\mathbb{S}_n$ is the unit sphere of $\mathbb{R}^n$ for the Euclidean norm $\|.\|$, whose inner product is $\langle , \rangle$.

Linear Algebra. Vectors of $\mathbb{R}^n$ will be row vectors denoted by bold lowercase letters. A (row) matrix is denoted by $[\mathbf{b}_1, \ldots, \mathbf{b}_n]$. We denote by $\mathcal{M}_{m,n}(\mathcal{R})$ the set of $m \times n$ matrices over a ring $\mathcal{R}$. The group of $n \times n$ invertible matrices with real coefficients will be denoted by $\mathcal{GL}_n(\mathbb{R})$ and $\mathcal{O}_n(\mathbb{R})$ will denote the subgroup of orthogonal matrices. The transpose of a matrix M will be denoted by M^t, and M^{-t} will mean the inverse of the transpose. For a set $\mathcal{S}$ of vectors in $\mathbb{R}^n$ and $M \in \mathcal{M}_{n,m}(\mathbb{R})$, $\mathcal{S} \cdot M$ denotes the set $\{\mathbf{s} \cdot M : \mathbf{s} \in \mathcal{S}\}$. We denote by I_n the $n \times n$ identity matrix.

Rounding. We denote by $\lceil x \rfloor$ the closest integer to x. Naturally, $\lceil \mathbf{b} \rfloor$ denotes the operation applied to all the coordinates of $\mathbf{b}$.

Distributions. If X is a random variable, we denote by $\mathbb{E}[X]$ its expectation. For any set S, we denote by $\mathcal{U}(S)$ the uniform distribution over S, when applicable. If $\mathcal{D}$ is a distribution over $\mathbb{R}^n$, its *covariance* is the $n \times n$ symmetric positive matrix $\mathrm{Cov}(\mathcal{D}) = \mathbb{E}_{\mathbf{x} \leftarrow \mathcal{D}}[\mathbf{x}^t \mathbf{x}]$. The notation $\mathcal{D} \oplus \mathcal{D}'$ denotes the convolution of two distributions, that is the distribution of $\mathbf{x} + \mathbf{y}$ where $\mathbf{x} \leftarrow \mathcal{D}$ and $\mathbf{y} \leftarrow \mathcal{D}'$ are sampled independently. Furthermore, we denote by $\mathcal{D} \cdot B$ the distribution of $\mathbf{x}B$ where $\mathbf{x} \leftarrow \mathcal{D}$.

Zonotopes and Parallelepipeds. A zonotope is the Minkowski sum of finitely many segments. Here, we use centered zonotopes: the *zonotope* spanned by an $m \times n$ row matrix $V = [\mathbf{v}_1, \ldots, \mathbf{v}_m]$ is the set $\mathcal{Z}(V) = \{\sum_{i=1}^{m} x_i \mathbf{v}_i, -1 \leq x_i \leq 1\}$. We denote by $\mathcal{D}_{\mathcal{Z}(V)}$ the convolution distribution over $\mathcal{Z}(V)$ obtained by picking independently each x_i uniformly at random from $[-1, 1]^n$: in other words, $\mathcal{D}_{\mathcal{Z}(V)} = \mathcal{U}([-1, 1]^n) \cdot V$, which in general is not the uniform distribution over $\mathcal{Z}(V)$. However, in the particular case $V \in \mathcal{GL}_n(\mathbb{R})$, $\mathcal{Z}(V)$ is simply the *parallelepiped* $\mathcal{P}(V)$ spanned by V, and $\mathcal{D}_{\mathcal{P}(V)}$ is equal to the uniform distribution over $\mathcal{P}(V)$.

Differentials. Let f be a function from $\mathbb{R}^n$ to $\mathbb{R}$. The *gradient* of f at $\mathbf{w} \in \mathbb{R}^n$ is denoted by $\nabla f(\mathbf{w}) = (\frac{\partial f}{\partial x_1}(\mathbf{w}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{w}))$. The *Hessian matrix* of f at $\mathbf{w} \in \mathbb{R}^n$ is denoted by $\mathrm{H}\, f(\mathbf{w}) = (\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{w}))_{1 \leq i,j \leq n}$.

Running Times. All given running times were measured using a 2.27-GHz Intel Xeon E5520 core.

Lattices. We refer to the survey [24] for a bibliography on lattices. In this paper, by the term lattice, we mean a full-rank discrete subgroup of $\mathbb{R}^n$. A non-empty set $L \subseteq \mathbb{R}^n$ is a lattice if and only if there exists $B = [\mathbf{b}_1, \ldots, \mathbf{b}_n] \in \mathcal{GL}_n(\mathbb{R})$ such that $L = \{\sum_{i=1}^{n} n_i \mathbf{b}_i \mid n_i \in \mathbb{Z}\}$. Any such B is called a basis of L, and the absolute value of its determinant is the lattice volume $\mathrm{vol}(L)$ of the lattice L. The *closest vector problem* (CVP) is the following: given a basis of $L \subseteq \mathbb{Z}^n$ and a target $\mathbf{t} \in \mathbb{Q}^n$, find a lattice vector $\mathbf{v} \in L$ minimizing the distance $\|\mathbf{v} - \mathbf{t}\|$. If d is the minimal distance, then approximating CVP to a factor k means finding $\mathbf{v} \in L$ such that $\|\mathbf{v} - \mathbf{t}\| \leq kd$. *Bounded Distance Decoding* (BDD) is a special case of CVP where the distance to the lattice is known to be small.

2.2 The GGH Signature Scheme

The GGH scheme [9] works with a lattice L in $\mathbb{Z}^n$. The secret key is a non-singular matrix $R \in \mathcal{M}_n(\mathbb{Z})$, with very short row vectors. Following [21], the public key is the Hermite normal form (HNF) of L. The messages are hashed onto a "large enough" subset of $\mathbb{Z}^n$, for instance a large hypercube. Let $\mathbf{m} \in \mathbb{Z}^n$

be the hash of the message to be signed. The signer applies Babai's round-off CVP approximation algorithm [3] to get a lattice vector close to $\mathbf{m}$:

$$\mathbf{s} = \lfloor \mathbf{m}R^{-1} \rceil R, \tag{1}$$

so that $\mathbf{s} - \mathbf{m} \in \frac{1}{2}\mathcal{P}(R)$. To verify the signature $\mathbf{s}$ of $\mathbf{m}$, one checks that $\mathbf{s} \in L$ using the public basis B, and that the distance $\|\mathbf{s} - \mathbf{m}\|$ is sufficiently small.

2.3 NTRUSign

Basic scheme. NTRUSign [11] is an instantiation of GGH using the compact lattices from NTRU encryption [14], which we briefly recall: we refer to [11,4] for more details. In the former NTRU standards [4] proposed to IEEE P1363.1 [17], $N = 251$ and $q = 128$. Let $\mathcal{R}$ be the ring $\mathbb{Z}[X]/(X^N - 1)$ whose multiplication is denoted by $*$. One computes $(f, g, F, G) \in \mathcal{R}^4$ such that $f * G - g * F = q$ in $\mathcal{R}$ and f is invertible mod q, where f and g have 0–1 coefficients (with a prescribed number of 1), while F and G have slightly larger coefficients, yet much smaller than q. This quadruplet is the NTRU secret key. Then the secret basis is the following $(2N) \times (2N)$ block-wise circulant matrix:

$$R = \begin{bmatrix} \mathcal{C}(f) & \mathcal{C}(g) \\ \mathcal{C}(F) & \mathcal{C}(G) \end{bmatrix} \text{ where } \mathcal{C}(a) \text{ denotes } \begin{bmatrix} a_0 & a_1 & \cdots & a_{N-1} \\ a_{N-1} & a_0 & \cdots & a_{N-2} \\ \vdots & \ddots & \ddots & \vdots \\ a_1 & \cdots & a_{N-1} & a_0 \end{bmatrix},$$

and f_i denotes the coefficient of X^i of the polynomial f. Thus, the lattice dimension is $n = 2N$. Due to the special structure of R, a single row of R is sufficient to recover the whole secret key. Because f is chosen invertible mod q, the polynomial $h = g/f \mod q$ is well-defined in $\mathcal{R}$: this is the NTRU public key. Its fundamental property is that $f * h \equiv g \mod q$ in $\mathcal{R}$. The polynomial h defines the following (natural) public basis of the lattice: $\begin{bmatrix} I_n & \mathcal{C}(h) \\ 0 & qI_n \end{bmatrix}$, which implies that the lattice volume is q^N.

The messages are assumed to be hashed in $\{0, \ldots, q-1\}^{2N}$. Let $\mathbf{m}$ be such a hash. We write $\mathbf{m} = (\mathbf{m}_1, \mathbf{m}_2)$ with $\mathbf{m}_i \in \{0, \ldots, q-1\}^N$. The signature is the vector $(\mathbf{s}, \mathbf{t}) \in \mathbb{Z}^{2N}$ which would have been obtained by applying Babai's round-off CVP approximation algorithm to $\mathbf{m}$, except that it is computed more efficiently using convolution products and can even be compressed (see [11]). We described the basic NTRUSign scheme [11], as used in half of the parameter choices of the former NTRU standards [4].

Perturbations. The second half of parameter choices of NTRU standards [4] use perturbation techniques [10,4,12] to strengthen security, which are described in Sect. 2.5. But there is a second change: instead of the standard NTRU secret key, one uses the so-called *transpose basis*, which is simply R^t, then the public basis remains the same, except that one defines the public key as $h = F/f = G/g \mod q$ rather than $h = g/f \mod q$.

New parameters. In the latest NTRU article [13], new parameters for NTRUSign have been proposed. These include different values of (N, q) and a different shape

for f and g: the coefficients of f and g are now in $\{0, \pm 1\}$, rather than $\{0, 1\}$ like in [11]. But the scheme itself has not changed.

2.4 The Nguyen-Regev Attack

We briefly recall the Nguyen-Regev attack [23], using a slightly different presentation. The NR attack solves the following idealized problem:

Problem 1 (The Hidden Parallelepiped Problem or HPP) *Let $V = [\mathbf{v}_1, \ldots, \mathbf{v}_n] \in \mathcal{GL}_n(\mathbb{R})$ and let $\mathcal{P}(V) = \{\sum_{i=1}^n x_i \mathbf{v}_i : x_i \in [-1, 1]\}$ be the parallelepiped spanned by V. The input to the HPP is a sequence of* $\mathrm{poly}(n)$ *independent samples from the uniform distribution $\mathcal{D}_{\mathcal{P}(V)}$. The goal is to find a good approximation of the rows of $\pm V$.*

In practice, instead of samples from $\mathcal{D}_{\mathcal{P}(V)}$, the attack uses $2(\mathbf{s} - \mathbf{m})$ for all given message-signature pairs $(\mathbf{m}, \mathbf{s})$: this distribution is heuristically close to $\mathcal{D}_{\mathcal{P}(V)}$ where R is the secret basis. To recover rows of R, the attack simply rounds the approximations found to integer vectors. The NR attack has two stages: morphing and minimization.

Morphing the Parallelepiped into a Hypercube. The first stage of the NR attack is to transform the hidden parallelepiped into a hidden hypercube (see Alg. 1), using a suitable linear transformation L. It is based on the following elementary lemma [23, Lemmas 1 and 2]:

Lemma 1. *Let $V \in \mathcal{GL}_n(\mathbb{R})$ and denote by $G \in \mathcal{GL}_n(\mathbb{R})$ the symmetric positive definite matrix $V^t V$. Then:*

- $\mathrm{Cov}(\mathcal{D}_{\mathcal{P}(V)}) = G/3$.
- *If $L \in \mathcal{GL}_n(\mathbb{R})$ satisfies $LL^t = G^{-1}$ and we let $C = VL$, then $C \in \mathcal{O}_n(\mathbb{R})$ and $\mathcal{D}_{\mathcal{P}(V)} \cdot L = \mathcal{D}_{\mathcal{P}(C)}$.*

Algorithm 1 Morphing$(\mathcal{X})$: Morphing a Parallelepiped into a Hybercube

Input: A set $\mathcal{X}$ of vectors $\mathbf{x} \in \mathbb{R}^n$ sampled from the uniform distribution $\mathcal{D}_{\mathcal{P}(V)}$ over a parallelepiped.

Output: A matrix L such that $\mathcal{D}_{\mathcal{P}(V)} \cdot L$ is close to $\mathcal{D}_{\mathcal{P}(C)}$ for some $C \in \mathcal{O}_n(R)$.

1: Compute an approximation G of $V^t V$ using the set $\mathcal{X}$, using $\mathrm{Cov}(\mathcal{D}_{\mathcal{P}(V)}) = V^t V/3$ (see Lemma 1).
2: Return L such that $LL^t = G^{-1}$

This stage is exactly (up to scaling) the classical preprocessing used in independent component analysis to make covariance equal to the identity matrix:

Lemma 2. *Let G be the covariance matrix of a distribution $\mathcal{D}$ over $\mathbb{R}^n$. If $L \in \mathcal{GL}_n(\mathbb{R})$ satisfies $LL^t = G^{-1}$, then $\mathrm{Cov}(\mathcal{D} \cdot L) = I_n$.*

Learning a Hypercube. The second stage of the NR attack is to solve the hidden hypercube problem, using minimization with a gradient descent (see Alg. 2). Nguyen and Regev [23] showed that for any $V \in \mathcal{O}_n(\mathbb{R})$, if $\mathcal{D}$ denotes the distribution $\mathcal{D}_{\mathcal{P}(V)}$:

- The function $\mathrm{mom}_{\mathcal{D},4}(\mathbf{w}) = \mathbb{E}_{\mathbf{x}\leftarrow\mathcal{D}}[\langle \mathbf{x}, \mathbf{w}\rangle^4]$ has exactly $2n$ local minima over the unit sphere $\mathbb{S}_n$, which are located at $\pm\mathbf{v}_1, \cdots, \pm\mathbf{v}_n$, and are global minima.
- It is possible to find all minima of $\mathrm{mom}_{\mathcal{D},4}(\cdot)$ over $\mathbb{S}_n$ in random polynomial time, using Alg. 2 with parameter $\delta = 3/4$, thanks to the nice shape of $\mathrm{mom}_{\mathcal{D},4}(\cdot)$. Alg. 2 is denoted by $\mathbf{Descent}(\mathcal{X}, \mathbf{w}, \delta)$ which, given a point $\mathbf{w} \in \mathbb{S}_n$, performs a suitable gradient descent using the sample set $\mathcal{X}$, and returns an approximation of some $\pm\mathbf{v}_i$.

Algorithm 2 Descent$(\mathcal{X}, \mathbf{w}, \delta)$: Solving the Hidden Hypercube Problem by Gradient Descent

Input: A set $\mathcal{X}$ of samples from the distribution $\mathcal{D}_{\mathcal{P}(V)}$ where $V \in \mathcal{O}_n(\mathbb{R})$, a vector $\mathbf{w}$ chosen uniformly at random from $\mathbb{S}_n$ and a descent parameter δ.
Output: An approximation of some row of $\pm V$.

1: Compute an approximation $\mathbf{g}$ of the gradient $\nabla \mathrm{mom}_{V,4}(\mathbf{w})$ using $\mathcal{X}$.
2: Let $\mathbf{w}_{new} = \mathbf{w} - \delta\mathbf{g}$.
3: Divide $\mathbf{w}_{new}$ by its Euclidean norm $\|\mathbf{w}_{new}\|$.
4: **if** $\mathrm{mom}_{V,4}(\mathbf{w}_{new}) \geq \mathrm{mom}_{V,4}(\mathbf{w})$ where the moments are approximated using $\mathcal{X}$ **then**
5: **return** the vector $\mathbf{w}$.
6: **else**
7: Replace $\mathbf{w}$ by $\mathbf{w}_{new}$ and go back to Step 1.
8: **end if**

The whole NR attack is summarized by Alg. 3.

Algorithm 3 SolveHPP$(\mathcal{X})$: Learning a Parallelepiped [23]

Input: A set $\mathcal{X}$ of vectors $\mathbf{x} \in \mathbb{R}^n$ sampled from $\mathcal{D}_{\mathcal{P}(V)}$, where $V \in \mathcal{GL}_n(\mathbb{R})$
Output: An approximation of a random row vector of $\pm V$

1: $L :=$ **Morphing**$(\mathcal{X})$ using Alg. 1
2: $\mathcal{X} := \mathcal{X} \cdot L$
3: Pick $\mathbf{w}$ uniformly at random from $\mathbb{S}_n$
4: Compute $\mathbf{r} :=$ **Descent**$(\mathcal{X}, \mathbf{w}, \delta) \in \mathbb{S}^n$ using Alg. 2: use $\delta = 3/4$ in theory and $\delta = 0.7$ in practice.
5: Return $\mathbf{r}L^{-1}$

Shrinking the number of NTRUSign*-signatures.* In practice, the NR attack requires a polynomial number of signatures, but it is possible to experimentally

decrease this amount by a linear factor [23], using a well-known symmetry of NTRU lattices. We define the NTRUSign symmetry group $\mathfrak{S}_N^{\mathrm{NTRU}}$ as the group spanned by $\sigma \in \mathcal{O}_n(\mathbb{R}) : (x_1, \dots x_N | y_1, \cdots y_N) \mapsto (x_2, \dots x_N, x_1 | y_2, \cdots y_N, y_1)$. If L is the NTRU lattice, then $\sigma(L) = L$. Furthermore, $(\sigma(\mathbf{m}), \sigma(\mathbf{s}))$ follows the same distribution as uniformly random $(\mathbf{m}, \mathbf{s})$. So, any pair $(\mathbf{m}, \mathbf{s})$ gives rise to N parallelepiped samples. This technique also allows a N-factor speedup for covariance computation, which is the most time consuming part of the attack.

2.5 Countermeasures

NTRUSign *perturbation: Summing Parallelepipeds.* Roughly speaking, these techniques perturbates the hashed message $\mathbf{m}$ before signing it with the NTRU secret basis. More precisely, the hashed message $\mathbf{m}$ is first signed using a second NTRU secret basis (of another NTRU lattice, which is kept secret), and the resulting signature is then signed as before. Heuristically, the effect on the sample distribution of the transcript is as follows: if R and R' are the two secret bases, the distribution of $\mathbf{s} - \mathbf{m}$ becomes the convolution $\mathcal{P}(R) \oplus \mathcal{P}(R')$, *i.e.* a natural distribution over the Minkowski sum of the two parallelepipeds obtained by adding the uniform distributions of both parallelepipeds.

IEEE-IT perturbation: Parallelepiped Deformation. Hu *et al.* [16] suggested another approach to secure NTRUSign in the journal IEEE Trans. IT. Their definition are specific to NTRUSign-bases, but it can be generalized to GGH, and we call this technique "Parallelepiped deformation". Let $\delta : [-1/2, 1/2)^n \to \mathbb{Z}^n$ be a function, possibly secret-key dependent. The signature generation (1) is replaced by:

$$\mathbf{s} = \Big(\lceil \mathbf{m} R^{-1} \rfloor + \delta\big(\mathbf{m} R^{-1} - \lceil \mathbf{m} R^{-1} \rfloor\big) \Big) R \tag{2}$$

If δ outputs small integer vectors, then the signature $\mathbf{s}$ is still valid. The associated deformation function is $d_\delta(\mathbf{x}) = \mathbf{x} + \delta(\mathbf{x})$. The sample distribution of $\mathbf{s} - \mathbf{m}$ is deformed in the following way : $d_\delta(\mathcal{U}^n) \cdot R$ where $d_\delta(\mathcal{U}^n)$ denotes the distribution of $\mathbf{x} + \delta(\mathbf{x})$ with $\mathbf{x} \leftarrow \mathcal{U}^n$. In [16], the deformation δ_{IEEE} for a NTRUSign secret key (f, g, F, G) is as follows:

- Let $U \subset [N]$ be the set of indexes u such that the u-th entry of $f + g + F + G$ is 1 modulo 2, and let $A = \#U$. On the average, $A \approx N/2$, and it is assumed that $A \geq 25$, otherwise a new secret key must be generated.
- Let $1 \leq u_1 < u_2 < \cdots < u_A \leq N$ be the elements of U. For $i \notin [A]$, u_i denotes $u_{(i \bmod A)}$.
- Let the input of δ_{IEEE} be the concatenation of two vectors $\mathbf{x}, \mathbf{y} \in [-1/2, 1/2)^N$. Then the i-th entry of $\delta_{\mathrm{IEEE}}(\mathbf{x}|\mathbf{y})$ is:

$$\big[\delta_{\mathrm{IEEE}}(\mathbf{x}|\mathbf{y})\big]_i = \begin{cases} 0 & \text{if } i \notin U \\ s(x_{u_j}, y_{u_j}, y_{u_{j+1}}, y_{u_{j+3}}, y_{u_{j+7}}, y_{u_{j+12}}) & \text{if } i = u_j \end{cases}$$

$$\text{where } s(a_0, \dots, a_5) = \begin{cases} 1 & \text{if } a_i < 0 \text{ for all } i \\ -1 & \text{if } a_i > 0 \text{ for all } i \\ 0 & \text{otherwise} \end{cases}$$

Gaussian Sampling. Gentry *et al.* [8] described the first provably secure countermeasure: Gaussian sampling. In previous schemes, the distribution of $\mathbf{s} - \mathbf{m}$ was related to the secret key. In [8], the distribution becomes independent of the secret key: it is some discrete Gaussian distribution, which gives rise a to a security proof in the random-oracle model, under the assumption that finding close vectors is hard in the NTRU lattice. Unfortunately, this countermeasure is not very competitive in practice: the sampling algorithm [18] is much less efficient than NTRUSign generation, and the new signature is less close to the message, which forces to increase parameters. But its efficiency has recently been improved, see [26,6].

3 Learning a Zonotope: Breaking NTRUSign with Perturbations

In Sect. 3.1, we introduce the hidden zonotope problem (HZP), which is a natural generalization of the hidden parallelepiped problem (HPP), required to break NTRUSign with perturbations. In Sect. 3.2, we explain why the Nguyen-Regev HPP algorithm (Alg. 3) can heuristically solve the HZP, in cases that include NTRUSign, provided that Step 5 is slightly modified. Yet, the approximations obtained by the algorithm are expected to be worse than in the non-perturbed case, so we use a folklore meet-in-the-middle algorithm for BDD in NTRU lattices, which is described in [5]. Finally, in Sect. 3.3, we present experimental results with our optimized NR attack which show that NTRUSign with one (or slightly more) perturbation(s) is completely insecure, independently of the type of basis. In particular, we completely break the original NTRUSign proposed to IEEE P1363 standardization [4]: only one half of the parameter sets was previously broken in [23].

3.1 The Hidden Zonotope Problem

Assume that one applies $k-1$ NTRUSign perturbations as a countermeasure, which corresponds to k NTRUSign lattices $L_1, \ldots, L_k$ (with secret bases $R_1, \ldots, R_k$) where only L_k is public. One signs a hashed message $\mathbf{m} \in \mathbb{Z}^n$ by computing $\mathbf{s}_1 \in L_1$ such that $\mathbf{s}_1 - \mathbf{m} \in \frac{1}{2}\mathcal{P}(R_1)$, then $\mathbf{s}_2 \in L_2$ such that $\mathbf{s}_2 - \mathbf{s}_1 \in \frac{1}{2}\mathcal{P}(R_2)$, ..., and finally $\mathbf{s}_k \in L_k$ such that $\mathbf{s}_k - \mathbf{s}_{k-1} \in \frac{1}{2}\mathcal{P}(R_k)$. It follows that $\mathbf{s}_k$ is somewhat close to $\mathbf{m}$, because $\mathbf{s}_k - \mathbf{m}$ is in the Minkowski sum $\frac{1}{2}\mathcal{P}(R_1) + \frac{1}{2}\mathcal{P}(R_2) + \cdots + \frac{1}{2}\mathcal{P}(R_k)$, which is a zonotope spanned by $\frac{1}{2}R_1, \ldots, \frac{1}{2}R_k$. And heuristically, the distribution of $2(\mathbf{s}_k - \mathbf{m})$ is the convolution of all the k uniform distributions $\mathcal{D}_{\mathcal{P}(R_i)}$. In other words, similarly to the perturbation-free case, an attacker wishing to recover the secret key of a GGH-type signature scheme using perturbations using a polynomial number of signatures is faced with the following problem with $m = kn$:

Problem 2 (The Hidden Zonotope Problem or HZP) *Let $m \geq n$ be integers, and $V = [\mathbf{v}_1, \ldots, \mathbf{v}_m]$ be an $m \times n$ row matrix of rank n. The input to the HZP is a sequence of* $\mathrm{poly}(n, m)$ *independent samples from $\mathcal{D} =$*

$\mathcal{D}_{\mathcal{Z}(V)}$ *over* $\mathbb{R}^n$*, which is the convolution distribution over the zonotope* $\mathcal{Z}(V) = \{\sum_{i=1}^m x_i \mathbf{v}_i, -1 \leq x_i \leq 1\}$ *spanned by* V*. The goal is to find a good approximation of the rows of* $\pm V$*.*

Here, we assume V to have rank n, because this is the setting of NTRUSign with perturbation, and because the HPP is simply the HZP with $m = n$.

3.2 Extending the Nguyen-Regev Analysis to Zonotopes

Here, we study the behavior of the original Nguyen-Regev algorithm for learning a parallelepiped (**SolveHPP**$(\mathcal{X})$, Alg. 3) on a HZP instance, that is, when the secret matrix V is not necessarily square, but is an arbitrary $m \times n$ matrix of rank n with $m \geq n$. To do this, we need to change the analysis of Nguyen and Regev [23], and we will have to slightly change Alg. 3 to make the attack still work: Alg. 4 is the new algorithm. Recall that the input distribution $\mathcal{D}_{\mathcal{Z}(V)}$ is formed by $\sum_{i=1}^m x_i \mathbf{v}_i$ where the x_i's are uniformly chosen in $[-1, 1]$. We study how the two stages of the NR attack behave for $\mathcal{D}_{\mathcal{Z}(V)}$.

Morphing Zonotopes. We start with a trivial adaptation of Lemma 1 to zonotopes:

Lemma 3. *Let* V *be an* $m \times n$ *matrix over* $\mathbb{R}$ *of rank* n*. Let* G *be the symmetric definite positive matrix* $V^t V$*. Then:*

- $\mathrm{Cov}(\mathcal{D}_{\mathcal{Z}(V)}) = G/3$.
- *If* $L \in \mathcal{GL}_n(\mathbb{R})$ *satisfies* $LL^t = G^{-1}$ *and we let* $C = VL$*, then* $C^t C = I_n$ *and* $\mathcal{D}_{\mathcal{Z}(V)} \cdot L = \mathcal{D}_{\mathcal{Z}(C)}$.

Lemma 3 shows that if we apply **Morphing**$(\mathcal{X})$ (Alg. 1) to samples from $\mathcal{D}_{\mathcal{Z}(V)}$ (rather than $\mathcal{D}_{\mathcal{P}(V)}$), the output transformation L will be such that $\mathcal{D}_{\mathcal{Z}(V)} \cdot L$ is close to $\mathcal{D}_{\mathcal{Z}(C)}$ for some $m \times n$ matrix C such that $C^t C = I_n$.

In other words, the effect of Step. 2 in **SolveHPP**$(\mathcal{X})$ (Alg. 3) is to make the zonotope matrix V have orthonormal columns: $V^t V = I_n$. The following lemma gives elementary properties of such matrices, which will be useful for our analysis:

Lemma 4. *Let* V *be an* $m \times n$ *row matrix* $[\mathbf{v}_1, \ldots, \mathbf{v}_m]$ *such that* $V^t V = I_n$*. Then:*

- $\|\mathbf{w}\|^2 = \sum_{i=1}^m \langle \mathbf{w}, \mathbf{v}_i \rangle^2$ *for all* $\mathbf{w} \in \mathbb{R}^n$.
- $\|\mathbf{v}_i\| \leq 1$ *for all* $1 \leq i \leq m$.
- $\sum_{i=1}^m \|\mathbf{v}_i\|^2 = n$ *and* $\mathrm{Exp}_{\mathbf{x} \leftarrow \mathcal{U}(\mathbb{S}_n)}(\|\mathbf{x} V V^t\|^2) = n/m$.

Learning an "Orthogonal" Zonotope. Nguyen and Regev [23] used the target function $\mathrm{mom}_{\mathcal{D},4}(\mathbf{w}) = \mathbb{E}_{\mathbf{x} \leftarrow \mathcal{D}}[\langle \mathbf{x}, \mathbf{w} \rangle^4]$ for $\mathbf{w} \in \mathbb{S}_n$, $\mathcal{D} = \mathcal{D}_{\mathcal{P}(V)}$ and $V \in \mathcal{O}_n(\mathbb{R})$ to recover the hidden hypercube. We need to study this function when $\mathcal{D}$ is the zonotope distribution $\mathcal{D} = \mathcal{D}_{\mathcal{Z}(V)}$ to recover the hidden zonotope. Nguyen and Regev [23] gave elementary formulas for $\mathrm{mom}_{\mathcal{D},4}$ and $\nabla \mathrm{mom}_{\mathcal{D},4}$ when $\mathcal{D} = \mathcal{D}_{\mathcal{P}(V)}$ and $V \in \mathcal{O}_n(\mathbb{R})$, which can easily be adapted to the zonotope distribution $\mathcal{D}_{\mathcal{Z}(V)}$ if $V^t V = I_n$, as follows:

Lemma 5. *Let V be a $m \times n$ matrix over $\mathbb{R}$ such that $V^tV = I_n$, and $\mathcal{D}$ be the convolution distribution $\mathcal{D}_{\mathcal{Z}(V)}$ over the zonotope spanned by V. Then, for any $\mathbf{w} \in \mathbb{R}^n$:*

$$\mathrm{mom}_{\mathcal{D},4}(\mathbf{w}) = \frac{1}{3}\|\mathbf{w}\|^4 - \frac{2}{15}\sum_{i=1}^{m}\langle \mathbf{v}_i, \mathbf{w}\rangle^4$$

$$\nabla \mathrm{mom}_{\mathcal{D},4}(\mathbf{w}) = \frac{4}{3}\mathbf{w} - \frac{8}{15}\sum_{i=1}^{m}\langle \mathbf{v}_i, \mathbf{w}\rangle^3 \mathbf{v}_i \quad \textit{if} \quad \mathbf{w} \in \mathbb{S}_n$$

Corollary 1. *Under the same hypotheses as Lemma 5, the minima over $\mathbb{S}_n$ of the function $\mathrm{mom}_{\mathcal{D},4}(\mathbf{w})$ are the maxima (over $\mathbb{S}_n$) of $f(\mathbf{w}) = \sum_{i=1}^{m} f_{\mathbf{v}_i}(\mathbf{w})$ where $f_{\mathbf{v}}(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w}\rangle^4$ is defined over $\mathbb{R}^n$.*

In [23, Lemma 3], Nguyen and Regev used Lagrange multipliers to show that when $V \in \mathcal{O}_n(\mathbb{R})$, the local minima of $\mathrm{mom}_{\mathcal{D}_{\mathcal{P}(V)},4}$ were located at $\pm\mathbf{v}_1, \ldots, \mathbf{v}_n$, and these minima are clearly global minima. However, this argument breaks down when V is a rectangular $m \times n$ matrix of rank n such that $V^tV = I_n$. To tackle the zonotope case, we use a different argument, which requires to study each function $f_{\mathbf{v}_i}(\mathbf{w}) = \langle \mathbf{v}_i, \mathbf{w}\rangle^4$ individually:

Lemma 6. *Let $\mathbf{v} \in \mathbb{R}^n$ and $f_{\mathbf{v}}(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w}\rangle^4$ for $\mathbf{w} \in \mathbb{R}^n$. Then:*

1. *The gradient and Hessian matrix of $f_{\mathbf{v}}$ are $\nabla f_{\mathbf{v}}(\mathbf{w}) = 4\langle \mathbf{w}, \mathbf{v}\rangle^3 \cdot \mathbf{v}$ and $\mathrm{H}\, f_{\mathbf{v}}(\mathbf{w}) = 12\langle \mathbf{w}, \mathbf{v}\rangle^2 \cdot \mathbf{v}^t\mathbf{v}$.*
2. *There are only two local maxima of $f_{\mathbf{v}}$ over $\mathbb{S}_n$, which are located at $\pm\mathbf{v}/\|\mathbf{v}\|$, and their value is $\|\mathbf{v}\|^4$.*
3. *The local minima of $f_{\mathbf{v}}$ over $\mathbb{S}_n$ are located on the hyperplane orthogonal to $\mathbf{v}$, and their value is 0.*
4. *The mean value of $f_{\mathbf{v}}$ over $\mathbb{S}_n$ is $3\|\mathbf{v}\|^4/(n(n+2))$.*

This already gives a different point of view from Nguyen and Regev in the special case where $V \in \mathcal{O}_n(\mathbb{R})$: for all $1 \leq j \leq n$, $\mathbf{v}_j$ is a local maximum of $f_{\mathbf{v}_j}$ and a local minimum of $f_{\mathbf{v}_i}$ for all $i \neq j$ because $\mathbf{v}_i \perp \mathbf{v}_j$; and therefore $\pm\mathbf{v}_1, \ldots, \mathbf{v}_n$ are local extrema of $\mathrm{mom}_{\mathcal{U}\cdot V,4}$.

In the general case where V is an $m \times n$ matrix such that $V^tV = I_n$, let $\mathbf{d}_i = \mathbf{v}_i/\|\mathbf{v}_i\| \in \mathbb{S}_n$ for $1 \leq i \leq m$. The direction $\mathbf{d}_j$ is a local maximum of $f_{\mathbf{v}_j}$ over $\mathbb{S}_n$. On the other hand, $f_{\mathbf{v}_i}(\mathbf{d}_j)$ is likely to be small for $i \neq j$. This suggests that $\mathbf{d}_j$ should be very close to a local maximum of the whole sum $\sum_{i=1}^{m} f_{\mathbf{v}_i}(\mathbf{d}_j)$, provided that the local maximum $\|\mathbf{v}_j\|^4$ of $f_{\mathbf{v}_j}$ is somewhat larger than $\sum_{i\neq j} f_{\mathbf{v}_i}(\mathbf{d}_j)$. In fact, this local maximum $\mathbf{d}_j$ is intuitively shifted by $\mathbf{g}/(2\|\mathbf{v}_j\|^4)$ where $\mathbf{g}$ is the gradient of $\sum_{i=1}^{m} f_{\mathbf{v}_i}(\mathbf{d}_j)$ at $\mathbf{d}_j$, because this is exactly what happens for its second-order Taylor approximation. This is formalized by our main result, which provides a sufficient condition on V guaranteeing that a given direction $\mathbf{v}_j/\|\mathbf{v}_j\|$ is close to a local minimum of $\mathrm{mom}_{\mathcal{D}_{\mathcal{Z}(V)},4}$:

Theorem 3 (Local Minima for Zonotopes). *Let V be a $m \times n$ matrix over $\mathbb{R}$ such that $V^tV = I_n$. Assume that there is $\alpha \geq 1$ such that V is α-weakly-orthogonal, that is, its m rows satisfy for all $i \neq j$: $|\langle \mathbf{v}_i, \mathbf{v}_j \rangle| \leq \alpha \|\mathbf{v}_i\| \|\mathbf{v}_j\| / \sqrt{n}$. Let $1 \leq j \leq m$ and $0 < \varepsilon < 1/\sqrt{2}$ such that:*

$$\varepsilon \|\mathbf{v}_j\|^4 > 6 \left(\frac{\alpha}{\sqrt{n}} + \varepsilon \right)^2 \varepsilon + \frac{4}{\|\mathbf{v}_j\|^3} \| \sum_{i \neq j} \langle \mathbf{v}_j, \mathbf{v}_i \rangle^3 \mathbf{v}_i \| \tag{3}$$

which holds in particular if $\|\mathbf{v}_j\| \geq \dfrac{2\sqrt{\alpha}}{n^{1/12}}$ and $\varepsilon = \dfrac{5\alpha^3}{\sqrt{n}\|\mathbf{v}_j\|^4} < 1/\sqrt{2}$. Then, over the unit sphere, the function $\mathrm{mom}_{\mathcal{D}_{\mathcal{Z}(V)},4}$ has a local minimum at some point $\mathbf{m}_j \in \mathbb{S}_n$ such that $\mathbf{m}_j$ is close to the direction of $\mathbf{v}_j$, namely:

$$\left\langle \mathbf{m}_j, \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|} \right\rangle > 1 - \frac{\epsilon^2}{2} \quad \textit{and} \quad \left\| \mathbf{m}_j - \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|} \right\| \leq \epsilon.$$

And the local minimum $\mathrm{mom}_{\mathcal{D}_{\mathcal{Z}(V)},4}(\mathbf{m}_j)$ discloses an approximation of $\|\mathbf{v}_j\|$, namely:

$$\left| \mathrm{mom}_{\mathcal{D}_{\mathcal{Z}(V)},4}(\mathbf{m}_j) - \left(\frac{1}{3} - \frac{2\|\mathbf{v}_j\|^4}{15} \right) \right| \leq \frac{2}{15} \left(5\varepsilon^3 + 6\varepsilon^2 + 4\varepsilon + m \left(\varepsilon + \frac{\alpha}{\sqrt{n}} \right)^4 \right).$$

Proof. (Sketch of the proof in [5]) Let $\mathcal{B} = \{\mathbf{w} \in \mathbb{S}_n : \|\mathbf{w} - \mathbf{d}_j\| < \varepsilon\}$ be the open ball of $\mathbb{S}_n$ of radius ε, where $\mathbf{d}_j = \mathbf{v}_j / \|\mathbf{v}_j\| \in \mathbb{S}_n$ Notice that for all $\mathbf{w} \in \mathbb{S}_n$:

$$\|\mathbf{w} - \mathbf{d}_j\|^2 = \|\mathbf{w}\|^2 + \|\mathbf{d}_j\|^2 - 2\langle \mathbf{w}, \mathbf{d}_j \rangle = 2(1 - \langle \mathbf{w}, \mathbf{d}_j \rangle).$$

Therefore $\mathcal{B} = \left\{ \mathbf{w} \in \mathbb{S}_n : \langle \mathbf{d}_j, \mathbf{w} \rangle > 1 - \varepsilon^2/2 \right\}$, whose closure and boundary are denoted respectively by $\bar{\mathcal{B}}$ and $\partial\mathcal{B}$. Recall that $f = \sum_{i=1}^m f_{\mathbf{v}_i}$. We will prove the following property:

$$\forall \mathbf{w} \in \partial\mathcal{B}, f(\mathbf{w}) < f(\mathbf{d}_j), \tag{4}$$

which allows to conclude the proof of Th. 3. Indeed, by continuity, the restriction of f to $\bar{\mathcal{B}}$ has a global maximum at some point $\mathbf{m}_j \in \bar{\mathcal{B}}$. And (4) implies that $\mathbf{m}_j \notin \partial\mathcal{B}$, therefore $\mathbf{m}_j \in \mathcal{B}$. Thus, $\mathbf{m}$ is a global maximum of f over the open set $\mathcal{B}$: in other words, $\mathbf{m}_j$ is a local maximum of f, and therefore a local minimum of $\mathrm{mom}_{\mathcal{D},4}$. Furthermore, by definition of $\mathcal{B}$, we have: $\|\mathbf{m_j} - \mathbf{d}_j\| < \varepsilon$ and $\langle \mathbf{d}_j, \mathbf{m}_j \rangle > 1 - \varepsilon^2/2$. And the final inequality follows from:

$$\mathrm{mom}_{\mathcal{D},4}(\mathbf{m}_j) - \left(\frac{1}{3} - \frac{2\|\mathbf{v}_j\|^4}{15} \right) = \frac{2}{15} \left(\langle \mathbf{v}_j, \mathbf{d}_j \rangle^4 - \langle \mathbf{v}_j, \mathbf{m}_j \rangle^4 - \sum_{i \neq j} \langle \mathbf{v}_i, \mathbf{m_j} \rangle^4 \right).$$

We now prove (4). Let $\mathbf{w} \in \partial\mathcal{B}$. To show $f(\mathbf{d}_j) - f(\mathbf{w}) > 0$, we decompose it as:

$$\left(f_{\mathbf{v}_j}(\mathbf{d}_j) - f_{\mathbf{v}_j}(\mathbf{w}) \right) + \sum_{i \neq j} \left(f_{\mathbf{v}_i}(\mathbf{d}_j) - f_{\mathbf{v}_i}(\mathbf{w}) \right) \tag{5}$$

On the one hand, the left-hand term of (5) is:

$$f_{\mathbf{v}_j}(\mathbf{d}_j) - f_{\mathbf{v}_j}(\mathbf{w}) = \|\mathbf{v}_j\|^4 - (1 - \frac{\varepsilon^2}{2})^4 \|\mathbf{v}_j\|^4 \geq \varepsilon^2 \|\mathbf{v}_j\|^4 \tag{6}$$

because $\varepsilon < 1/\sqrt{2}$. On the other hand, we upper bound the right-hand term of (5) by the Taylor-Lagrange formula, which states that there exists $\theta \in (0,1)$ such that $\sum_{i \neq j} (f_{\mathbf{v}_i}(\mathbf{w}) - f_{\mathbf{v}_i}(\mathbf{d}_j))$ is equal to:

$$\left\langle \sum_{i \neq j} \nabla f_{\mathbf{v}_i}(\mathbf{d}_j), \mathbf{w} - \mathbf{d}_j \right\rangle + \frac{1}{2}(\mathbf{w} - \mathbf{d}_j) \sum_{i \neq j} \mathrm{H}\, f_{\mathbf{v}_i}(\mathbf{d}_j + \theta(\mathbf{w} - \mathbf{d}_j))(\mathbf{w} - \mathbf{d}_j)^t \tag{7}$$

Let $\mathbf{g} = \sum_{i \neq j} \nabla f_{\mathbf{v}_i}(\mathbf{d}_j) = 4 \sum_{i \neq j} \langle \mathbf{d}_j, \mathbf{v}_i \rangle^3 \mathbf{v}_i$ by Lemma 6. The left-hand term of (7) is bounded as:

$$\left| \left\langle \sum_{i \neq j} \nabla f_{\mathbf{v}_i}(\mathbf{d}_j), \mathbf{w} - \mathbf{d}_j \right\rangle \right| \leq \varepsilon \|\mathbf{g}\|. \tag{8}$$

Using Lemma 6, the right-hand term of (7) can be bounded as:

$$\left| (\mathbf{w} - \mathbf{d}_j) \sum_{i \neq j} \mathrm{H}\, f_{\mathbf{v}_i}(\mathbf{d}_j + \theta(\mathbf{w} - \mathbf{d}_j))(\mathbf{w} - \mathbf{d}_j)^t \right| \leq 12(\alpha/\sqrt{n} + \varepsilon)^2 \varepsilon^2. \tag{9}$$

Collecting (6), (7), (8) and (9), we obtain:

$$f(\mathbf{d}_j) - f(\mathbf{w}) \geq \left(\varepsilon \|\mathbf{v}_j\|^4 - \|\mathbf{g}\| - 6(\alpha/\sqrt{n} + \varepsilon)^2 \varepsilon \right) \varepsilon,$$

which is > 0 by (3). To conclude, it remains to prove that (3) is satisfied when $\|\mathbf{v}_j\| \geq \frac{2\sqrt{\alpha}}{n^{1/12}}$ and $\varepsilon = \frac{5\alpha^3}{\sqrt{n}\|\mathbf{v}_j\|^4} < 1/\sqrt{2}$. This is shown by tedious computations, using weak-orthogonality and Lemma 4. □

Th. 3 states that under suitable assumptions on V (which we will discuss shortly), if $\|\mathbf{v}_j\|$ is not too small, then the secret direction $\mathbf{v}_j/\|\mathbf{v}_j\|$ is very close to a local minimum of $\mathrm{mom}_{\mathcal{D}_{\mathcal{Z}(V)},4}$, whose value discloses an approximation of $\|\mathbf{v}_j\|$, because it is $\approx \frac{1}{3} - \frac{2}{15}\|\mathbf{v}_j\|^4$. This suggests **SolveHZP**$(\mathcal{X})$ (Alg. 4) for learning a zonotope: **SolveHZP**$(\mathcal{X})$ is exactly **SolveHPP**$(\mathcal{X})$ (Alg. 3), except that Step 5 of **SolveHPP**$(\mathcal{X})$ has been modified, to take into account that $\|\mathbf{v}_j\|$ is no longer necessarily equal to 1, but can fortunately be approximated by the value of the local minimum.

First, we discuss the value of α in Th. 3 . Note that weak-orthogonality is a natural property, as shown by the following basic result:

Lemma 7. *Let $\mathbf{v} \in \mathbb{S}^n$ and denote by X the random variable $X = \langle \mathbf{v}, \mathbf{w} \rangle^2$ where $\mathbf{w}$ has uniform distribution over $\mathbb{S}_n$. Then X has distribution Beta$(1/2, (n-1)/2)$, $\mathrm{Exp}(X) = \frac{1}{n}$, $\mathrm{Exp}(X^2) = \frac{3}{n(n+2)}$, $\mathrm{Exp}(X^3) = \frac{15}{n(n+2)(n+4)}$ and more generally: $\mathrm{Exp}(X^k) = \frac{k-1/2}{n/2+k-1}\mathrm{Exp}(X^{k-1})$.*

Algorithm 4 SolveHZP($\mathcal{X}$): Learning a Zonotope

Input: A set $\mathcal{X}$ of vectors $\mathbf{x} \in \mathbb{R}^n$ sampled from $\mathcal{D}_{\mathcal{Z}(V)}$, where V is an $m \times n$ matrix of rank n.

Output: An approximation of some row vector of $\pm V$.

1: $L := \textbf{Morphing}(\mathcal{X})$ using Alg. 1
2: $\mathcal{X} := \mathcal{X} \cdot L$
3: Pick $\mathbf{w}$ uniformly at random from $\mathbb{S}_n$
4: Compute $\mathbf{r} := \textbf{Descent}(\mathcal{X}, \mathbf{w}, \delta) \in \mathbb{S}^n$ using Alg. 2: use $\delta = 3/4$ in theory and $\delta = 0.7$ in practice.
5: Return $\lambda \mathbf{r} L^{-1}$ where $\lambda = ((\frac{1}{3} - \mathrm{mom}_{\mathcal{X},4}(\mathbf{r}))\frac{15}{2})^{1/4}$

By studying more carefully the Beta distribution, it is possible to obtain strong bounds. For instance, Ajtai [1, Lemma 47] showed that for all sufficiently large n, if $\mathbf{v} \in \mathbb{S}^n$ is fixed and $\mathbf{w}$ has uniform distribution over $\mathbb{S}_n$, then $|\langle \mathbf{v}, \mathbf{w} \rangle| \leq (\log n)/\sqrt{n}$ with probability $\geq 1 - \frac{1}{n^{(\log n)/2-1}}$. Since the probability is subexponentially close to 1, this implies that if $m = n^{O(1)}$ and we assume that all the directions $\mathbf{v}_i/\|\mathbf{v}_i\|$ are random, then V is $(\log n)$-weakly orthogonal with probability asymptotically close to 1.

This gives strong evidence that, if $m = n^{O(1)}$, the assumption on V in Th. 3 will be satisfied for $\alpha = \log n$. We can now discuss the remaining assumptions. If $\alpha = \log n$, we may take any index j such that $\|\mathbf{v}_j\| \geq \Omega(1/n^{13})$: in particular, if $\|\mathbf{v}_j\| = \Omega(1)$, we may take $\varepsilon = O(\log^3 n)/\sqrt{n}$. And higher values of α can be tolerated, as while as $\alpha = o(n^{1/6})$. Now recall that $\sum_{i=1}^m \|\mathbf{v}_i\|^2 = n$, thus $\max_i \|\mathbf{v}_i\| \geq \sqrt{n/m}$ and $\|\mathbf{v}_i\|$ is on average $\sqrt{n/m}$. In particular, if the number of perturbations is constant, then $m = O(n)$ and $\max_i \|\mathbf{v}_i\| \geq \Omega(1)$, therefore Th. 3 applies to at least one index j, provided that $\alpha = o(n^{1/6})$. In fact, one can see that the result can even tolerate slightly bigger values of m than $\Theta(n)$, such as $m = o(n^{7/6}/\log n)$.

While Th. 3 explains why **SolveHZP**($\mathcal{X}$) (Alg. 4) can heuristically solve the HZP, it is not a full proof, as opposed to the simpler parallelepiped case. The obstructions are the following:

- First, we would need to prove that the distance is sufficiently small to enable the recovery of the original zonotope vectors, using an appropriate BDD solver. Any error on $\mathbf{v}_j/\|\mathbf{v}_j\|$ is multiplied by $L^{-1}\|\mathbf{v}_j\|$. In [23], the error on $\mathbf{v}_j$ could be made polynomially small for any polynomial, provided that the number of samples was (polynomially) large enough. But ε cannot be chosen polynomially small for any arbitrary polynomial in Th. 3.
- Second, we would need to prove that **Descent**($\mathcal{X}, \mathbf{w}, \delta$) (Alg. 2) finds a random local minimum of $\mathrm{mom}_{\mathcal{D}_{\mathcal{Z}(V)},4}$ in polynomial time, even in the presence of noise to compute $\mathrm{mom}_{\mathcal{D}_{\mathcal{Z}(V)},4}$. Intuitively, this is not unreasonable since the function $\mathrm{mom}_{\mathcal{D}_{\mathcal{Z}(V)},4}$ is very regular, but it remains to be proved.
- Finally, we would need to prove that there are no other local minima, or at least, not too many of them.

Regarding the third obstruction, it is easy to prove the following weaker statement, which implies that global minima of $\mathrm{mom}_{\mathcal{D}_{\mathcal{Z}(V)},4}$ over the unit sphere are close to some direction $\mathbf{v}_j/\|\mathbf{v}_j\|$:

Lemma 8. *Let V be a $m \times n$ matrix over $\mathbb{R}$ such that $V^tV = I_n$, and $\mathcal{D}$ be the distribution $\mathcal{D}_{\mathcal{Z}(V)}$. Let $\mathbf{w}$ be a global maximum of $f(\mathbf{w}) = \sum_{i=1}^{m} f_{\mathbf{v}_i}(\mathbf{w})$ over $\mathbb{S}_n$. Then there exists $j \in \{1, \ldots, m\}$ such that: $\frac{1}{m^{1/4}} < \frac{|\langle \mathbf{v}_j, \mathbf{w} \rangle|}{\|\mathbf{v}_j\|} \leq 1$.*

3.3 Experiments

We now report on experiments with the attack performed on NTRUSign, with n up to 502. Our experiments are real-world experiments using signatures of uniformly distributed messages.

Conditions of Th. 3. Our discussion following Th. 3 suggested that the matrix V should be heuristically weakly-orthogonal for $\alpha = \log n$. In practice, we may in fact take $\alpha \approx 5$ for both types of NTRUSign secret bases.

Regarding the norms $\|\mathbf{v}_i\|$ after morphing, we experimentally verified that $\|\mathbf{v}_i\| \approx \sqrt{1/k}$ where k is the number of perturbations for NTRUSign transposed bases (see [5]), as expected by $\sum_{i=1}^{m} \|\mathbf{v}_i\|^2 = n$. But for the so-called standard bases, the situation is a bit different: half of the $\|\mathbf{v}_i\|$'s are very small, and the remaining half are close to $\sqrt{2/k}$. This can be explained by the fact that standard bases are unbalanced: half of the vectors are much shorter than the other vectors.

For a number of perturbations ≤ 8, we experimentally verified that the "gradient" $\mathbf{g} = \frac{4}{\|\mathbf{v}_j\|^3} \| \sum_{i \neq j} \langle \mathbf{v}_j, \mathbf{v}_i \rangle^3 \mathbf{v}_i \|$ appearing in the conditions of Th. 3 satisfies $\|\mathbf{g}\| = O(1/n)$ with a small constant ≤ 4 (see [5]).

To summarize, the conditions of Th. 3 are experimentally verified for a number of perturbations ≤ 8: for all vectors $\mathbf{v}_j$'s in the case of transposed bases, and for half of the vectors $\mathbf{v}_j$'s in the case of standard bases.

Modifications to the original NR attack. We already explained that the original NR algorithm **SolveHPP**($\mathcal{X}$) (Alg. 3) had to be slightly modified into **SolveHZP**($\mathcal{X}$) (Alg. 4): more precisely, Step 5 is modified.

However, because Th. 3 states that the secret direction might be perturbed by some small ε, we also implemented an additional modification: instead of the elementary BDD algorithm by rounding, we used in the final stage a special BDD algorithm tailored for NTRU lattices, which is a tweaked version of Odlyzko's meet-in-the-middle attack on NTRU described in [15]. Details are given in [5].

Practical cryptanalysis. We first applied successfully the optimized NR-attack on the original NTRUSign-251 scheme with one perturbation (which corresponds to a lattice dimension of 502), as initially submitted to the IEEE P1363 standard: about 8,000 signatures were sufficient to recover the secret key, which should be compared with the 400 signatures of the original attack [23] when there was no

perturbation. This means that the original NTRUSIGN-251 scheme [10] is now completely broken.

Furthermore, we performed additional experiments for varying dimension and number of perturbations, for the parameters proposed in the latest NTRU article [13], where transposed bases are used. Table 1 summarizes the results obtained: each successful attack took less than a day, and the MiM error recovery algorithm ran with less than 8Gb of memory.

Table 1. Experiments with the generalized NR-attack on the latest NTRUSIGN parameters [13]

Security level : dimension n	Toy : 94	80-bit : 314	112-bit : 394	128-bit : 446
0 perturbation	300:(0,1)	400:(0,1)	400:(0,1)	600:(0,1)
1 perturbation	1000:(1,2)	5000:(0,1)	4000:(0,1)	4000:(0,0)
2 perturbations	10000:(5,3)	12000:(0,2)		
3 perturbations	12000:(5,4)			
4 perturbations	100000:(0,1)			

In this table, each non-empty cell represents a successful attack for a given transposed basis (the column indicates the security level and the dimension) and number of perturbations (row). These cells have the form $s : (e = \|\epsilon_F\|_1, w = \|\epsilon_G\|_\infty)$ where s is the number of signatures used by the learning algorithm, and where $(\epsilon_F|\epsilon_G)$ is the error vector of the best approximation given by a descent. The running time of our MiM-Algorithm is about $(n/2)^{\lceil e/2 \rceil + 1}$ for such small w.

Our experiments confirm our theoretical analysis: NTRUSIGN with a constant number of perturbations is insecure, but we see that the number of signatures required increases with the number of perturbations.

4 Learning a Deformed Parallelepiped: Breaking the IEEE-IT Countermeasure

In this section, we show that the deformation suggested in [16] is unlikely to prevent the NR attack [23]. More generally, we show that the NR attack heuristically still works if the deformation is only *partial*, which means that it preserves at least one of the canonical axes, namely there exists at least one index i such that:

- for all $\mathbf{x} \in [-1/2, 1/2)^n$, $[\delta(\mathbf{x})]_i = 0$
- $\delta(\mathbf{x})$ is independent of $x_i : (\forall j \neq i, x_j = y_j) \Rightarrow \delta(\mathbf{x}) = \delta(\mathbf{y})$

Such an index i is said to be ignored by the deformation δ. And it is clear that δ_{IEEE} is partial by definition (see Sect. 2.5), because it ignores exactly all index $i \notin U$. Our main result is the following, whose proof is given in [5].

Theorem 4. *Let δ be a partial deformation, and i be an index ignored by δ. Let $\mathcal{D} = 2 \cdot d_\delta(\mathcal{U}^n)$ and $M \in \mathcal{GL}_n(\mathbb{R})$ be an invertible matrix and $G = \text{Cov}(\mathcal{D} \cdot M)$. Let*

L be such that $LL^t = G^{-1}$. Then $\mathbf{r} = \frac{1}{\sqrt{3}} \cdot \mathbf{m}_i L$ is a local minimum of $\mathrm{mom}_{4,\mathcal{D}'}(\cdot)$ over the unit sphere, where $\mathcal{D}' = \mathcal{D} \cdot M \cdot L$.

While this is a strong theoretical argument supporting why the NR attack still works, it is not a full proof, for reasons similar to the zonotope case (see the previous section): there may be other minima, and we did not prove that the gradient descent efficiently finds minima.

Experimental results The attack was run, using 300,000 signatures, to recover the secret key in 80-bit, 112-bit and 128-bit NTRUSign security level settings, and each run led to a secret key recovery, in about two days. No other local minimum was found. Though the samples no longer belong to a set stable by NTRU symmetry group $\mathfrak{S}_N^{\mathrm{NTRU}}$, we may still try to apply the symmetry trick, to multiply the number of samples by N, like in [23]. This modifies the distribution of the sample to the average of its orbit : $\mathfrak{S}_N^{\mathrm{NTRU}}(\mathcal{D}) = \sigma(\mathbf{x}) : \mathbf{x} \leftarrow \mathcal{D}, \sigma \leftarrow \mathcal{U}(\mathfrak{S}_N^{\mathrm{NTRU}})$. It turns out that applying the attack on such an averaged distribution leads once again to descents converging to some basis vectors: in fact, by symmetry, all of them are equally likely. The attack used 2,000 signatures, and ran in less than an hour, on the same basis. Intuitively, this averaging strongly reduces the co-dependence between the coordinates of $\mathbf{x} \leftarrow \mathcal{D}_\sigma$, making the resulting distribution much closer to a parallelepiped than $\mathcal{D}$.

Acknowledgements. Part of this work is supported by the Commission of the European Communities through the ICT program under contract ICT-2007-216676 ECRYPT II, the European Research Council, and by China's 973 Program (Grant 2013CB834205).

References

1. M. Ajtai. Generating random lattices according to the invariant distribution. Draft of March 2006.
2. M. Ajtai. Generating hard instances of lattice problems. In *Complexity of computations and proofs*, volume 13 of *Quad. Mat.*, pages 1–32. Dept. Math., Seconda Univ. Napoli, Caserta, 2004.
3. L. Babai. On Lovász lattice reduction and the nearest lattice point problem. *Combinatorica*, 6:1–13, 1986.
4. Consortium for Efficient Embedded Security. Efficient embedded security standards #1: Implementation aspects of NTRUEncrypt and NTRUSign. Version 2.0 available at [17], June 2003.
5. L. Ducas and P. Q. Nguyen. Learning a zonotope and more: Cryptanalysis of NTRUSign countermeasures. Full version of the ASIACRYPT '12 article.
6. L. Ducas and P. Q. Nguyen. Faster Gaussian lattice sampling using lazy floating-point arithmetic. In *Proc. ASIACRYPT '12*, Lecture Notes in Computer Science. Springer, 2012.
7. C. Gentry. Fully homomorphic encryption using ideal lattices. In *Proc. STOC '09*, pages 169–178. ACM, 2009.

8. C. Gentry, C. Peikert, and V. Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *Proc. STOC '08*, pages 197–206. ACM, 2008.
9. O. Goldreich, S. Goldwasser, and S. Halevi. Public-key cryptosystems from lattice reduction problems. In *Proc. of Crypto '97*, volume 1294 of *LNCS*, pages 112–131. IACR, Springer-Verlag, 1997. Full version vailable at ECCC as TR96-056.
10. J. Hoffstein, N. A. H. Graham, J. Pipher, J. H. Silverman, and W. Whyte. NTRUSIGN: Digital signatures using the NTRU lattice. Full version of [11]. Draft of April 2, 2002, available on NTRU's website.
11. J. Hoffstein, N. A. H. Graham, J. Pipher, J. H. Silverman, and W. Whyte. NTRUSIGN: Digital signatures using the NTRU lattice. In *Proc. of CT-RSA*, volume 2612 of *LNCS*. Springer-Verlag, 2003.
12. J. Hoffstein, N. A. H. Graham, J. Pipher, J. H. Silverman, and W. Whyte. Performances improvements and a baseline parameter generation algorithm for NTRUsign. In *Proc. of Workshop on Mathematical Problems and Techniques in Cryptology*, pages 99–126. CRM, 2005.
13. J. Hoffstein, N. Howgrave-Graham, J. Pipher, and W. Whyte. Practical lattice-based cryptography: NTRUEncrypt and NTRUSign. 2010. In [25].
14. J. Hoffstein, J. Pipher, and J. Silverman. NTRU: a ring based public key cryptosystem. In *Proc. of ANTS III*, volume 1423 of *LNCS*, pages 267–288. Springer-Verlag, 1998. First presented at the rump session of Crypto '96.
15. N. Howgrave-Graham, J. H. Silverman, and W. Whyte. A meet-in-the-middle attack on an NTRU private key. In *http://www.ntru.com/cryptolab/tech_notes.htm#004*, 2003.
16. Y. Hu, B. Wang, and W. He. NTRUSign with a new perturbation. *IEEE Transactions on Information Theory*, 54(7):3216–3221, 2008.
17. IEEE P1363.1. Public-key cryptographic techniques based on hard problems over lattices. See `http://grouper.ieee.org/groups/1363/lattPK/index.html`, June 2003.
18. P. Klein. Finding the closest lattice vector when it's unusually close. In *Proc. of SODA '00*. ACM–SIAM, 2000.
19. V. Lyubashevsky. Lattice signatures without trapdoors. *IACR Cryptology ePrint Archive*, 2011:537, 2011. In EUROCRYPT '12.
20. T. Malkin, C. Peikert, R. A. Servedio, and A. Wan. Learning an overcomplete basis: Analysis of lattice-based signatures with perturbations. 2009 manuscript cited in [26], available as [27, Chapter 6].
21. D. Micciancio. Improving lattice-based cryptosystems using the Hermite normal form. In *Proc. of CALC '01*, volume 2146 of *LNCS*. Springer-Verlag, 2001.
22. D. Micciancio and O. Regev. Lattice-based cryptography. In *Post-quantum cryptography*, pages 147–191. Springer, Berlin, 2009.
23. P. Q. Nguyen and O. Regev. Learning a Parallelepiped: Cryptanalysis of GGH and NTRU Signatures. *J. Cryptology*, 22(2):139–160, 2009. Preliminary version in EUROCRYPT 2006.
24. P. Q. Nguyen and J. Stern. The two faces of lattices in cryptology. In *Proc. of CALC '01*, volume 2146 of *LNCS*. Springer-Verlag, 2001.
25. P. Q. Nguyen and B. Vallée, editors. *The LLL Algorithm: Survey and Applications*. Information Security and Cryptography. Springer, 2010.
26. C. Peikert. An efficient and parallel gaussian sampler for lattices. In *Proc. CRYPTO '10*, volume 6223 of *Lecture Notes in Computer Science*, pages 80–97. Springer, 2010.
27. A. Wan. *Learning, cryptography, and the average case*. PhD thesis, Columbia University, 2010. Available at `http://itcs.tsinghua.edu.cn/~atw12/`.