NTRU Cryptosystem: Recent Developments and Emerging Mathematical Problems in Finite Polynomial Rings

Ron Steinfeld

Abstract. The NTRU public-key cryptosystem, proposed in 1996 by Hoffstein, Pipher and Silverman, is a fast and practical alternative to classical schemes based on factorization or discrete logarithms. In contrast to the latter schemes, it offers quasi-optimal asymptotic efficiency and conjectured security against quantum computing attacks. The scheme is defined over finite polynomial rings, and its security analysis involves the study of natural statistical and computational problems defined over these rings.

We survey several recent developments in both the security analysis and in the applications of NTRU and its variants, within the broader field of lattice-based cryptography. These developments include a provable relation between the security of NTRU and the computational hardness of worst-case instances of certain lattice problems, and the construction of fully homomorphic and multilinear cryptographic algorithms. In the process, we identify the underlying statistical and computational problems in finite rings.

Keywords. NTRU Cryptosystem, lattice-based cryptography, fully homomorphic encryption, multilinear maps.

AMS classification. 68Q17, 68Q87, 68Q12, 11T55, 11T71, 11T22.

1 Introduction

The NTRU public-key cryptosystem has attracted much attention by the cryptographic community since its introduction in 1996 by Hoffstein, Pipher and Silverman [32, 33]. Unlike more classical public-key cryptosystems based on the hardness of integer factorisation or the discrete logarithm over finite fields and elliptic curves, NTRU is based on the hardness of finding 'small' solutions to systems of linear equations over polynomial rings, and as a consequence is closely related to geometric problems on certain classes of high-dimensional Euclidean *lattices*. From a practical point of view, the distinguishing feature of NTRU compared with classical systems, has mainly been its very high *speed* of encryption and decryption operations for practical security levels under best known attacks, being faster than classical systems by 2 or more orders of magnitude. This highly attractive feature has led to the inclusion of NTRU in the

The author was supported by an Australian Research Fellowship (ARF) from the Australian Research Council (ARC), and Discovery Grants DP0987734 and DP110100628.

IEEE P1363 industry standard for cryptography [36]. It is also often considered as the most viable 'post-quantum' public-key encryption due to its conjectured resistance to attack by quantum computers (see, e.g., [58]), whereas classical systems have been shown [70] to be insecure in the presence of quantum computing.

In this survey, we focus on recent exciting developments in both the security analysis and applications of NTRU, that we believe should make the NTRU system even more attractive for study and development in future than the above 'traditional' reasons, and suggest new motivation and directions for studying NTRU's mathematical underpinnings.

In terms of security of NTRU, one of the troubling issues since its introduction, has been a lack of confidence in the hardness of its underlying computational problems. We review the computational/statistical problems underlying NTRU's security, and a recent result of Stehlé and Steinfeld [71, 72], that shows a variant of NTRU whose security can be proved based on the *worst-case quantum hardness* of natural lattice problems over the class of *ideal* lattices defined over certain polynomial rings. Although the hardness of the latter problem is not guaranteed, the result of [71, 72] shows that any efficient attack against the NTRU variant implies a significant advance in computational algebraic number theory: an efficient quantum algorithm for the problem of finding an element of small Euclidean norm in *any* given ideal of the underlying ring. In the process of describing this result, we introduce some Fourier analysis tools which find common use today in the wider field of lattice-based cryptography.

In terms of applications of NTRU, we review two recent novel variants of the NTRU system, which allow powerful new functionality to be added to the basic cryptosystem. The first application is an NTRU-based Fully Homomorphic Encryption (FHE) scheme [43], which allows useful computation on encrypted messages, and the second is a construction for NTRU-based multilinear maps [23] (as simplified by [40]), which open the door to another class of applications including non-interactive multiparty key agreement.

There are interesting recent developments related to NTRU that we do *not* cover due to space limitations, and we only mention some of them here. The hard problems underlying the NTRU cryptosystem can also be used to design a digital signature scheme. There is a long 'design-break-repair' history behind this scheme, now known as NTRUSign (see, e.g., [26, 29, 35, 74, 53, 54] and the survey [31]). Recent notable developments in this area include a variant of NTRUSign with a security proof based on worst-case hardness of lattice problems [72], and an alternative particulary efficient class of signature schemes [44, 45, 21], the most recent of which [21] is also based on NTRU-like hardness assumptions, that can be considered a lattice analogue of Schnorr's discrete-log based signature scheme [69]. Another recent line of investigation is generalizations of NTRU, that use higher degree algebraic rings to replace the integer ground ring in NTRU (see, e.g. [37]). We do not attempt in this work to survey the whole area and history of lattice-based cryptography, but refer the interested reader instead to the surveys [51, 62] for some starting points.

The rest of this paper is organized as follows. In Section 2, we introduce some notation used throughout. After a review of the basic NTRU cryptosystem and its underlying mathematical problems in Section 3, we survey some central tools in the Stehlé-Steinfeld security proof for NTRU in Section 4. Beginning with a review of the underlying Fourier analysis tools in Section 4.2, we explain their application to analyzing the 'statistical region' NTRU key cracking problem in Section 4.3, and then in Section 4.4 briefly look at how the NTRU ciphertext cracking problem relates to the now well-known Ring-LWE problem and worst-case lattice problems. Moving to recent novel applications of NTRU in Section 5, in Section 5.1 we review the homomorphic cryptosystem of López-Alt et al. [43] and its underlying computational problems, and in Section 5.2 we explain the multilinear maps of Garg et al. [23] and underlying problems, ending with some concluding remarks in Section 6.

2 Notation and Preliminaries

Notation. We make use of the Landau notations $O(\cdot), o(\cdot), \omega(\cdot), \Omega(\cdot)$. A function $f(n)$ is called *negligible* if $f(n) = n^{-\omega(1)}$, otherwise, if there exists a constant $c > 0$ such that $f(n) > n^{-c}$ for infinitely many n, we say that $f(n)$ is *non-negligible* and write $f(n) = \text{non-neg}(n)$. We write $\text{poly}(n)$ to denote a polynomial function of n. We say that a sequence of events E_n holds with overwhelming probability if $\Pr[\neg E_n] \leq f(n)$ for a negligible function f.

Probability and Algorithms. A probabilistic algorithm is called *efficient* if its running time is polynomial in its input length (in the cryptographic settings discussed in this survey, the input length is always polynomial in a *security parameter* n, and so one can think of running time and probabilities as all being functions of n). We write $x \hookleftarrow D$ to denote that x is a random variable sampled from the probability distribution D. The effectiveness of an algorithm to distinguish between two probability distributions D_0 and D_1 is measured by its *distinguishing advantage*, defined by $|\Pr_{x \hookleftarrow D_0}[A(x) = 1] - \Pr_{x \hookleftarrow D_1}[A(x) = 1]|$. We say that a decision problem (parameterized by n) is *hard* if there does not exist an efficient algorithm for it that has a non-negligible advantage in n. The *statistical distance* $\Delta(D_0; D_1)$ between two distributions D_0, D_1 on some countable domain X is defined as $\Delta(D_0; D_1) \stackrel{\text{def}}{=} \frac{1}{2} \cdot \sum_{x \in X} |D_0(x) - D_1(x)|$, and is extended to a continuous domain X and density functions ν_0 and ν_1 on X via $\Delta(\nu_0; \nu_1) \stackrel{\text{def}}{=} \frac{1}{2} \cdot \int_X |\nu_0(x) - \nu_1(x)| dx$. We denote by $U(X)$ the uniform distribution on a finite discrete domain X, or the uniform density function on X if X is finite and continuous.

Rings. If q is a non-zero integer, we let $\mathbb{Z}_q$ denote the ring of integers modulo q, i.e., the set $\{0, \ldots, q-1\}$ with addition and multiplication modulo q. We sometimes identify $\mathbb{Z}_q$ with the set of residues reduced into the interval $(-q/2, q/2]$. For a ring $(R, +, \times)$, we let R^* denote the set of invertible elements of R. For the polynomial ring $R = \mathbb{Z}[x]/(x^n + 1)$, if $x, y \in R$, their product $x \cdot y$ is the polynomial whose

coefficient vector is given by the vector-matrix product $y^T \cdot \mathrm{rot}(x)$, where $y^T \in \mathbb{Z}^n$ denotes the coefficient row vector of y and $\mathrm{rot}(x)$ is the nega-cyclic $n \times n$ matrix with entries in $\mathbb{Z}$, having the coefficient vector of x on its first row and each additional row is obtained from the one above it by rotating one position to the right and negating the leftmost entry. If $x \in R$ is an element of a polynomial ring R, we denote by $\|x\|$ the Euclidean norm of the coefficient vector of x.

Lattices. A (full-rank) *lattice* is a set of the form $L = \sum_{i \leq n} \mathbb{Z} \cdot b_i$, where the b_i's are linearly independent vectors in $\mathbb{R}^n$. The integer n is called the *lattice dimension*, and the b_i's are called a *basis* of L. The *minimum* $\lambda_1(L)$ (resp. $\lambda_1^\infty(L)$) is the Euclidean (resp. infinity) norm of any shortest non-zero vector of L. More generally, we define the k-th *successive minimum* $\lambda_k(L)$ for any $k \leq n$ as the smallest r such that L contains at least k linearly independent vectors of norm $\leq r$. The *dual lattice* of L is defined as $\widehat{L} = \{c \in \mathbb{R}^n : \forall i, \langle c, b_i \rangle \in \mathbb{Z}\}$. If $B = (b_i)_i$ is a basis matrix of L, the *fundamental parallelepiped* of B is the set $\mathcal{P}(B) = \{\sum_{i \leq n} c_i b_i : c_i \in [0, 1)\}$. The volume $|\det B|$ of $\mathcal{P}(B)$ is an invariant of the lattice L, denoted by $\det L$. The most famous algorithmic problem on lattices is the Shortest Vector Problem (SVP). Given a basis of a lattice L, it aims at finding a shortest vector in $L \setminus \{0\}$. It can be relaxed to an approximate-SVP (approx-SVP) variant γ-SVP by asking for a non-zero vector that is no longer than $\gamma(n)$ times a solution to SVP, for a prescribed function $\gamma(\cdot)$. It is believed that no subexponential quantum algorithm solves γ-SVP in the worst case, for any γ that is polynomial in the dimension. The smallest γ which is known to be achievable in polynomial time is exponential, up to poly-logarithmic factors in the exponent [41, 68, 52].

3 Review of the NTRU Cryptosystem

3.1 The NTRU Construction

We review the construction of the basic NTRU Public-Key Cryptosystem. We present it here in a general form, to allow us to easily explain later the difference in instantiation choices between the original cryptosystem [33] and more recent variants [71, 72].

The NTRU system is defined over a polynomial ring $R = \mathbb{Z}[x]/(\phi(x))$, for some polynomial $\phi(x) \in \mathbb{Z}[x]$ of degree n (a typical setting in the original system would take $\phi(x) = x^n - 1$ with $n = 257$). It is determined by a set of parameters $(n, q, p, \chi_\sigma, \chi_\rho, \chi_\beta)$, having the following interpretation. The parameter n is the degree of the modulus polynomial ϕ defining the polynomial ring R. Polynomials $p \in R$ and $q \in R$ are used to define the quotient polynomial rings $R_q = R/(qR)$ and $R_p = R/(pR)$ that form the ciphertext space and message space of the cryptosystem, respectively (A typical setting would take $q = 64$ and $p = 3$ as small integers). Strictly speaking, the message space should be considered a subset S_p of the ring R that has a unique representation modulo p. In particular, when p is a rational integer, R_p has p^n elements and if p is odd, the message space S_p can be taken as the set $S_p = \{-(p-1)/2, \ldots, (p-1)/2\}^n$ of n-

dimensional vectors with integer coordinates of magnitude $\leq (p-1)/2$ (note however, that in some settings, such as the multilinear map variant we discuss in Sec. 5.2, or as an efficiency optimization [34], it is useful to choose p as a polynomial of non-zero degree). Finally, $(\chi_\sigma, \chi_\rho, \chi_\beta)$ are probability distributions for the secret key polynomials, ciphertext secret polynomial, and ciphertext noise polynomial, respectively. They are defined over the ring R and, for correctness of the decryption algorithm, are concentrated on elements of *small* norm, compared to the modulus q. The distribution parameters σ, ρ, β are a measure of the norm of elements sampled from the respective distribution (typically, in the original NTRU system, χ_σ, χ_ρ outputs random polynomials with most coefficients being zero and the rest in the set $\{-1, 1\}$, while χ_β outputs the zero polynomial). Also, for correctness of decryption, the support of the distribution χ_σ is additionally restricted to elements $f \in R$ satisfying $f \bmod p = 1$ and $f \bmod q \in R_q^*$.

The key generation, encryption, and decryption algorithms are then defined as in Figure 1.

- **Key generation** $\mathsf{KG}(n)$**.** On input security parameter n:
  - Sample f and g from χ_σ.
  - Return secret key $f \in R_q^*$ and public key $h = g/f \in R_q^*$.
- **Encryption** $\mathsf{Enc}(h, m)$**.** Given public-key h and message $m \in S_p$, sample s from χ_ρ, sample e from χ_β, and return ciphertext $c = p \cdot (hs + e) + m \in R_q$.
- **Decryption** $\mathsf{Dec}(f, c)$**.** Given secret-key f and ciphertext c, compute $c' = f \cdot c \in R_q$ and return message $m = c' \bmod p$.

Figure 1. The basic NTRU scheme. Note that $e = 0$ and the public key is $p \cdot h$ in the original NTRU system.

To generate a secret/public key-pair for himself, a user Bob samples two polynomials of small norm f, g. Bob keeps the 'small' polynomial f as his secret key, and publishes the quotient $h = g/f \in R_q^*$ as his public key. Observe that although g, f have a small norm compared to q, their quotient mod q is typically 'large' mod q. The underlying intuition is that the public key h "looks" like a uniformly distributed element of R_q, to any potential eavesdropping adversary Cathy (we will come back to make this intuition more precise later on).

To encrypt a private message m intended for Bob, a user Alice takes Bob's public key h, samples 'small' polynomials s and e, computes the ciphertext $c = p \cdot (hs + e) + m \in R_q$ and sends it to Bob along a public communication channel (possibly eavesdropped by Cathy). The underlying intuition is that the term $p \cdot (hs + e)$ "looks" random to Cathy and "masks" the added message m in R_q (we will come back to this intuition later also). On the other hand, c does *not* look random to Bob, due to his knowledge of the small polynomial f. Indeed, Bob can decrypt c by computing

$c' = f \cdot c = p(gs + fe) + fm \in R_q$. Since p, g, s, f, e are all elements of R of small norm compared with q, both the terms $p(gs+fe)$ and fm are also of small norm compared with q, the coefficients of $p(gs + fe) + fm$ can be guaranteed (at least with high probability) to be all smaller than q, so that the equality $c' = p(gs + fe) + fm$ holds in R (and not just in R_q). As a consequence, Bob can recover the message m by computing c' mod $p = fm$ mod $p = m$ mod p, using f mod $p = 1$, and the fact that m mod p determines $m \in S_p$ uniquely by the unique representation of S_p modulo p. We remark here that in the original NTRU system [33], there was no error term ($e = 0$), and the public key was defined as $h = p \cdot g/f$, to save a multiplication by p during encryption. This 'original variant' does not achieve semantic security (as defined below), a problem that was addressed in [71] by introducing a non-zero error term e.

Setting the Parameters for Correct Decryption. To give an idea of parameter settings for NTRU to keep in mind for the rest of the survey, we give an example parameter setting as a function of the security parameter n, for getting perfectly correct decryption (in practice, less conservative parameters can often be used to improve efficiency, while tolerating a small decryption error probability). Let us assume that $\phi(x) = x^n - 1$ (as in the original NTRU proposal), $p = 3 = O(1)$ and suppose that (σ, ρ, β) are upper bounds on the infinity norm of ring elements sampled by $\chi_\sigma, \chi_\rho, \chi_\beta$ respectively. Since the ith coefficient of the polynomial $gs \in R$ is the inner-product between the coefficient vector of s and the coefficient vector of $x^i \cdot g(x)$ (the latter being just the coefficient vector of g rotated by i positions), we have by the Schwartz inequality $\|gs\|_\infty \leq \|g\| \cdot \|s\| \leq n \cdot \sigma\rho$. Similarly, we have $\|fe\|_\infty \leq n \cdot \sigma\beta$ and $\|fe\|_\infty \leq p\sigma$. The decryption correctness condition $\|p(gs + fe) + fm\|_\infty < q/2$ then is satisfied if $\sigma \cdot pn \cdot (\beta + \rho + 1) < q/2$. To satisfy the latter, one may then take for example $\sigma, \beta, \rho = O(1)$ and $q = O(n)$. In the next section, we will see that to allow a security proof for the scheme assuming the hardness of worst-case ideal lattice problems, one needs somewhat larger parameters, though still polynomial in n.

3.2 Security of NTRU: Computational/Statistical Problems and Known Attacks

Semantic Security of Public-Key Encryption Schemes. We will focus in this survey on the standard security notion for public-key encryption schemes against passive eavesdropping attacks, known as semantic security, or indistinguishability against chosen-plaintext attack. We refer here to this notion as IND, but it is often called IND-CPA in the cryptographic literature (see [39] for a good introduction to modern security notions for public-key encryption schemes, including more advanced active attacks that we do not consider). It asks that there exists no efficient algorithm A that has success probability $1/2 + \text{non-neg}(n)$ in the following two-phase 'game' between a challenger and the adversary A:

(i) The challenger gives A the public key pk sampled by the key generation algo-

rithm $\mathsf{KG}(n)$, and A returns two challenge messages m_0, m_1.

(ii) The challenger chooses a uniformly distributed bit $b \in \{0,1\}$ and gives A a challenge ciphertext $c_b = \mathsf{Enc}(pk, m_b)$ for message m_b.

(iii) A outputs an estimate b' for the bit b. We say that A succeeds if $b' = b$.

This notion of security, first defined by Goldwasser and Micali [30], guarantees that no 'partial' information on the message leaks to the adversary via the ciphertext. As noticed in [71], the original NTRU system with a zero error term e is *not* secure in the sense of IND. Indeed, given $c = hs + m_b$ and m_0, m_1, to compute b, one can compute the element $s_{b'} = (c - m_{b'}) \cdot h^{-1}$ in R_q for $b' \in \{0,1\}$ and return the value of b' such that $s_{b'}$ is 'small' compared to q (indeed, notice that $s_b = s$ is 'small', whereas $s_{1-b} = s + (m_b - m_{1-b})h^{-1}$ is very likely to be 'large', since h^{-1} is 'large'). This problem was fixed by introducing a non-zero error term e in the variant of NTRU proposed in [71] (we remark in passing that 'plain' variants of other cryptosystems, such as 'plain' RSA [63], are also insecure in the sense of IND, and there exist techniques to deal with the problem – see, e.g. [39]).

Computational Problems Arising from NTRU. There are essentially two ways to break the IND security of NTRU. The first is to recover the secret 'small' polynomials (f, g) from the public key h, or alternatively some other pair $(f', g') \in R^2$ of 'small' polynomials satisfying $h = g'/f' \bmod q$ (it is easy to see that as long as the norm $\|(f', g')\| = \ell$ is sufficiently small compared with q, the polynomials (f', g') can serve as an 'equivalent' secret key, in the sense that f' will also successfully decrypt NTRU ciphertexts), and use this to distinguish the challenge c – we call this approach the 'NTRU key cracking problem'. The second way is to distinguish the message encrypted in the ciphertext c "without using the special structure" of the public-key h – we call this approach the 'NTRU ciphertext cracking problem'.

In the following we define both a search and a decision version of the key cracking problem.

Definition 3.1 (NTRU Key Cracking Problems). The NTRU search and decision key cracking problems are defined as follows:

- Search Key Cracking Problem $\mathtt{NKC}_{n,p,q,\phi,\chi_\sigma,\ell}$: Given $h = g/f \in R_q$, with f, g sampled independently from χ_σ, compute $(f', g') \in R^2$ satisfying $h = g'/f' \bmod q$ and $\|(f', g')\| \leq \ell$.
- Decision Key Cracking Problem $\mathtt{DNKC}_{n,p,q,\phi,\chi_\sigma}$: Given $h \in R_q^*$, distinguish whether h is sampled from the distribution $D_0 = \{h = g/f \in R_q : f, g \hookleftarrow \chi_\sigma\}$ or from the distribution $D_1 = U(R_q^*)$.

For the ciphertext cracking problem, we can make the requirement "without using the special structure" precise by requiring that an algorithm for the 'NTRU ciphertext inversion' problem works even for a public key h that is uniformly distributed in R_q^*. This leads to the following definitions.

Definition 3.2 (NTRU Ciphertext Cracking Problems). The NTRU search and decision key cracking problems are defined as follows:

- Search Ciphertext Cracking Problem $\mathtt{NCC}_{n,p,q,\phi,\chi_\rho,\chi_\beta}$: Given h sampled from $U(R_q^*)$, and $c = p \cdot (hs + e) + m \in R_q$ with s, e sampled independently from χ_ρ, χ_β respectively, compute (s, e, m).
- Decision Ciphertext Cracking Problem $\mathtt{DNCC}_{n,p,q,\phi,\chi_\rho,\chi_\beta}$: Given h sampled from $U(R_q^*)$, and $c \in R_q$, distinguish whether c is sampled from the distribution $D_0 = \{c = p \cdot (hs + e) : s \hookleftarrow \chi_\rho, e \hookleftarrow \chi_\beta\}$ or from the uniform distribution $U(R_q)$.

The utility of the decision problems defined above is that we can actually prove that NTRU achieves semantic security if both problems are hard.

Proposition 3.3. *If the decision key cracking problem $\mathtt{DNKC}_{n,p,q,\phi,\chi_\sigma}$ and the decision ciphertext cracking problem $\mathtt{DNCC}_{n,p,q,\phi,\chi_\rho,\chi_\beta}$ are both hard, then the NTRU$_{n,p,q,\phi,\chi_\sigma,\chi_\rho,\chi_\beta}$ cryptosystem achieves semantic (IND) security.*

Proof. We argue by contradiction. Suppose A denotes an efficient algorithm for breaking the IND security of NTRU, having success probability $p = \Pr_{IND}[b' = b] = 1/2 + \text{non-neg}(n)$. Consider the modified IND game IND', in which the public-key h provided to A by the challenger is sampled from $U(R_q^*)$, instead of being generated by the key generation algorithm KG. Let $p' = \Pr_{IND'}[b' = b]$ denote the success probability of A in game IND'. If $|p' - p| = \text{non-neg}(n)$, then we can use A as an efficient distinguisher against the decision key cracking problem $\mathtt{DNKC}_{n,p,q,\phi,\chi_\sigma}$ with non-negligible advantage $|p' - p|$, a contradiction with the hardness of this problem. Indeed, such a distinguisher B, on input an instance $h \in R_q$ of the $\mathtt{DNKC}_{n,p,q,\phi,\chi_\sigma}$ problem, would run A with public-key input h, with B simulating the IND game by acting as a challenger for A. When A outputs b', B outputs 1 if $b' = b$ and 0 else. Observe that if B's input h comes from the key generation distribution $D_0 = \{h = g/f \in R_q : f, g \hookleftarrow \chi_\sigma\}$, B perfectly simulates for A the IND game, and hence B outputs 1 with probability p, whereas if B's input h comes from the uniform distribution $D_1 = U(R_q^*)$, B perfectly simulates for A the IND' game, and hence B outputs 1 with probability p'. It follows that B's distinguishing advantage is $|p' - p|$, as required, which leads by contradiction to the conclusion $p' = 1/2 + \text{non-neg}(n)$.

Now consider a further modified game IND" that is obtained from IND' by changing the distribution of the challenge ciphertext c_b to be sampled independently from $U(R_q)$. Let $p'' = \Pr_{IND''}[b' = b]$ denote the success probability of A in game IND". Because c_b is statistically independent of b in game IND", we have $p'' = 1/2$, which implies that $|p'' - p'| = \text{non-neg}(n)$. But this means (by a similar construction as used for distinguisher B above) that we can use A as an efficient distinguisher against the decision ciphertext cracking problem $\mathtt{DNCC}_{n,p,q,\phi,\chi_\rho,\chi_\beta}$ with non-negligible advantage $|p'' - p'|$, a contradiction with the hardness of this problem. □

Although not directly sufficient for the IND security proof of NTRU, the hardness of the search variants of the NTRU cracking problems is also clearly a necessary condition for the security of the decision problems, and forms a good starting point for the analysis of NTRU. We also remark that the hardness of the decision key cracking problem is *not* known to be necessary for the IND security of NTRU.

Known Attacks: link to lattices. The NTRU system was originally presented as a system based on hard problems in polynomial rings. However, soon after its presentation, Coppersmith and Shamir [19] realized that the NTRU cracking problems can actually be interpreted as computational problems on a related special class of lattices, now known as the *NTRU lattices*. By using existing algorithms for lattice reduction, in particular the LLL algorithm [41] and its many variants, lattice methods quickly became the dominant avenue of attacking the NTRU system. Other known attacks on the system are combinatorial in nature and easily avoided by a careful choice of parameters, so we will not cover them here (see [31] for details on these). Instead, we here explain the link to lattices that will lead to the more recent developments described in the next section.

The main observation of Coppersmith and Shamir was that the search key cracking problem $\mathtt{NKC}_{n,p,q,\phi,\chi_\sigma,\ell}$ can be interpreted as a variant of the shortest vector problem in a related lattice. Namely, the secret key $(f, g) \in R^2$ satisfies the homogenous linear relation

$$f \cdot h - g = 0 \bmod q. \tag{3.1}$$

in R. Hence, the set of all solutions $(f, g) \in R^2$ to (3.1) forms an R-module M_h over the ring R, and can be generated by the rows of the matrix

$$\begin{bmatrix} 1 & h \\ 0 & q \end{bmatrix}. \tag{3.2}$$

Recall (see Section 2) that when we represent the elements of R by their coefficient vectors in $\mathbb{Z}^n$, a ring multiplication $x \cdot y$ corresponds to a vector-matrix product $x^T \cdot \cdot \mathrm{rot}(y)$. It follows that the module M_h can also be viewed as a $2n$-dimensional lattice $L_{NTRU} \subseteq \mathbb{Z}^{2n}$ with row basis matrix

$$\begin{bmatrix} I_n & \mathrm{rot}(h) \\ 0 & qI_n \end{bmatrix}, \tag{3.3}$$

which is known as the *NTRU lattice*. The secret key (f, g) is a typically a "short" vector in L_{NTRU}, and thus approx-SVP algorithms can be used for computing (f, g) or a vector $v = (f', g') \in L_{NTRU}$ of not much larger norm. Although this method gives a practical algorithm for the key cracking problem for small n, it does not seem to give an efficient attack for large n. Indeed, to be useful, the norm ℓ of the non-zero vector v returned by the approx-SVP algorithm must be smaller than q. This is because, as we recall, successful decryption of a ciphertext $c = p \cdot (hs + e) + m$ using $v = (f', g')$ by

computing $f'c \bmod q = p(g's + f'e) + f'm \bmod q$, relies on $\|v\|$ being sufficiently small compared to q so that $\|p(g's+f'e)+f'm\|_\infty < q/2$ and hence $(f'c \bmod q) \bmod p = f'm \bmod p$ (for the same reason, the non-zero vector $(q, 0, \ldots, 0)$ of norm q, which is always in L_{NTRU}, is not useful for decryption of NTRU ciphertexts). On the other hand, recall that with typical parameters settings we have $q = \text{poly}(n)$, whereas $\|v\| \geq \lambda_1(L_{NTRU}) \geq 1$ since $L_{NTRU} \subseteq \mathbb{Z}^{2n}$. We conclude that, to be successful in attacking this problem, we need an approx-SVP algorithm with approximation factor $\gamma(n) < q = \text{poly}(n)$, and this is currently believed to require time exponential in n, at least for *worst-case* lattice instances.

However, the class of NTRU lattices in scope looks very special. This naturally suggests an open question that has clouded the understanding of NTRU's security ever since, namely: *Is the* poly(n)*-SVP problem restricted to the class of NTRU lattices as hard as the worst-case for general lattices? Is there a non-negligible fraction of NTRU lattices for which the* poly(n)*-SVP problem is easy?*

The complexity of known algorithms for the the approx-SVP problem for arbitrary lattices, namely the LLL algorithm and its variants, seems to behave the same for NTRU lattices as for other lattices, and the NTRU designers conjectured [33] that the approx-SVP for NTRU lattices is indeed hard. But the possibility still remains that a special approx-SVP algorithm may be tailored to exploit the structure of NTRU lattices, or a large fraction thereof, possibly by exploiting algebraic properties. This was a main motivation for the results we survey in the following section.

4 Recent Developments in Security Analysis of NTRU

4.1 Overview

Motivated by the uncertainty over the security of NTRU, Stehlé and Steinfeld [71, 72] studied the possibility of applying the tools developed over the last decade within the recently developing field of lattice-based cryptography, to obtain a variant of NTRU with a *provable* security guarantee with respect to the hardness of *worst-case* instances in a natural class of *ideal lattices*, corresponding to the ideals in the ring R, against quantum attacks. Although this result does not exclude the possibility that poly(n)-SVP is easy for random NTRU lattices, it does show that the realization of such a possibility would imply a significant development in computational algebraic number theory. In this Section, we explain the key ideas used to obtain the result of [71], and in the process introduce some powerful tools from lattice based cryptography.

The result of [71, 72] builds on a large body of work on the design of cryptographic schemes with provable security guarantees based on the worst-case hardness of lattice problems. Interestingly, this research field began independently around the same time the original NTRU system was introduced, with the pioneering work of Ajtai [1]. This work, followed by that of Ajtai-Dwork [2] and its later significant extension by Regev [59, 60], Micciancio-Regev [50], and Gentry et al. [27], laid the foundations

for this area by establishing tools for constructing (somewhat inefficient) one-way functions and public-key cryptosystems with security provably based on the (sometimes quantum) worst-case hardness of lattice problems over *general* lattices. Inspired by the high efficiency of the (then still heuristic) algebraic NTRU system, Micciancio [49] was the first to study adaptations of Ajtai's result to the design of *efficient* cryptographic one-way functions with a provable security guarantee, based on worst-case hardness of lattice problems restricted to classes of algebraically structured lattices, now known as *ideal lattices*. This work was further generalized and extended by Peikert-Rosen [57] and Lyubashevsky-Micciancio [46]. Later, concurrent and independent work by Stehle, Steinfeld, Tanaka, and Xagawa [73] and Lyubashevsky, Peikert and Regev [47, 48] adapted the results of Regev [60] to establish tools for designing efficient public-key encryption schemes based on worst-case (quantum) hardness of ideal lattice problems. In particular, the decision Ring-LWE problem introduced in [47, 48] and shown there to be as hard to solve as certain worst-case ideal lattice problems, was one of the main ingredients used in the NTRU security proof of [71] (see Sec. 4.4). The second main ingredient, introduced in [71], are statistical properties of NTRU-like lattices, building on statistical tools introduced in [50] and extended in [27] for general lattices (see Sec. 4.3). We have chosen to emphasize in this part of our survey the statistical tools of the second ingredient and their application to the NTRU key cracking problem, and we only briefly mention the computational methods of the first ingredient, since they are already well covered in Regev's survey [62] on the Learning With Errors (LWE) problem and its Ring-LWE variant.

We begin by stating the main result of Stehlé-Steinfeld [71]. To do so, we need to first define the class of worst-case lattices involved. Let $R = \mathbb{Z}[x]/\phi(x)$. By mapping polynomials to the vectors of their coefficients, we see that a non-zero ideal I of R corresponds to a full-rank sublattice of $\mathbb{Z}^n$: we can thus view I as both a lattice and an ideal. An *ideal lattice* for ϕ is a sublattice of $\mathbb{Z}^n$ that corresponds to a non-zero ideal $I \subseteq \mathbb{Z}[x]/\phi$. We say that A is an algorithm for the γ-Ideal-SVP$_\phi$ problem if, given as input the basis of *any* ideal lattice for ϕ (i.e. even 'worst-case' instances), it outputs a non-zero vector in the lattice of norm $\leq \gamma$ times the norm of the shortest non-zero vector in the lattice. The following result shows that there exists a choice of parameters for NTRU for which breaking the IND security of random instances is as hard as the worst-case hardness of poly(n)-SVP in ideal lattices.

Theorem 4.1 ([71]). *Fix $\varepsilon > 0$ and suppose n is a power of 2 such that $\phi(x) = x^n + 1$ splits into n linear factors modulo a prime $q = \Omega(n^{4.5+\varepsilon}\|p\|^4)$, for some $p \in R$ with $\deg p \leq 1$, and $R = \mathbb{Z}[x]/(\phi(x))$. There exist efficiently sampleable distributions $\chi_\sigma, \chi_\rho, \chi_\beta$ with norm parameters $\sigma = n\sqrt{\ln(8nq)} \cdot q^{1/2+\varepsilon}$ and $\rho = \beta$ with $q/\rho = \Omega(n^{0.25+\varepsilon})\|p\|^2\sigma$ and $\rho = \Omega(n^{1+\varepsilon})$ such that the following holds: If there exists an algorithm for breaking the IND security of NTRU$_{n,p,q,\chi_\sigma,\chi_\rho,\chi_\beta}$ which runs in time* poly(n) *and has success probability* $1/2 + 1/\text{poly}(n)$*, then there exists a* poly(n)*-time quantum algorithm for γ-*Ideal-SVP$_\phi$ *with $\gamma = O(n^{2.75+\varepsilon})\|p\|^2 q^{1/2+\varepsilon}$.*

Moreover, the decryption algorithm succeeds with probability $1 - n^{-\omega(1)}$.

Before we discuss the proof of Theorem 4.1, we make some remarks on the choice of parameters compared to those of the original NTRU system. The most significant change is the use of a wider deviation parameter $\sigma > q^{1/2}$ for the secret polynomials f, g (compared to f, g with ternary coefficients in the original system) – this is used to make the decision key cracking problem statistically hard. The non-zero distribution χ_β for the error term e (which is zero in the original system) is introduced to make the decision ciphertext cracking problem hard (as explained above). Another two technical changes have been made: the modulus polynomial has been changed to $\phi(x) = x^n + 1$ (rather than $\phi(x) = x^n - 1$ in the original scheme) due to its irreducibility over $\mathbb{Q}$, and the modulus q was chosen to satisfy the condition $q = 1 \bmod 2n$, which implies that $x^n + 1$ splits completely modulo q. These properties (assumed in the rest of this section) are needed for the computational reduction discussed in Section 4.4, between the Ring-LWE problem and worst-case lattice problems.

By applying Proposition 3.3, the proof of IND security of NTRU in Theorem 4.1 decomposes into two parts. The first part in Section 4.3 studies the security of the Decision Key Cracking Problem DNKC, while the second part in Section 4.4 looks at the Decision Ciphertext Cracking Problem DNCC. Recall that the decision key cracking problem asks to distinguish the distribution D_0 of the quotient $h = g/f \in R_q^*$ from the uniform distribution D_1 on R_q^*, when f and g are sampled from the distribution χ_σ that is concentrated on a subset S of R consisting of elements with coefficients of magnitude $\leq \sigma$. Since there are $|S \times S| \approx \sigma^{2n}$ pairs (f, g) in $S \times S$, there are two natural regions for this problem, depending on the value of σ, that we call the 'statistical region' and the 'computational region'.

In the 'statistical' region, we have $|S \times S| > |R_q^*| \approx q^n$ (or $\sigma > q^{1/2}$), and we heuristically expect that the distribution of h would 'fill' R_q^* approximately uniformly, i.e. that the distribution D_0 would be close to the uniform one on R_q^*. In the following, we explain the proof from [71] that shows that this is indeed the case when $\sigma > \text{poly}(n)q^{1/2}$, and the shape of the distribution χ_σ is a (discrete) Gaussian (restricted to R_q^*). Since the statistical distance between the distributions can be made negligibly small, this shows the *statistical hardness* of the decision key cracking problem $\text{DNKC}_{\chi_\sigma}$ for $\sigma \geq \text{poly}(n)q^{1/2}$. By 'statistical hardness', we mean that in this region, even adversaries with unlimited computational power cannot distinguish the decision key cracking problem with non-negligible advantage.

In the 'computational region', we have $|S \times S| < |R_q^*| \approx q^n$ (or $\sigma < q^{1/2}$). In this case the size of $|S \times S|$ is not sufficiently large to 'fill' the space R_q^*, and the distribution of h must be far from uniform on R_q^*. In this case, there always exists a distinguisher that achieves a non-negligible advantage between D_0 and D_1, and the best we can hope for is that there does not exist an *efficient* distinguisher, i.e. that it is *computationally hard* to distinguish D_0 from D_1. It is an interesting open problem to prove the computational hardness of this problem under a worst-case hardness assumption, such as the

worst-case hardness of approx-SVP in ideal lattices. In fact, little seems to be currently known even about the statistical properties of the distribution of h in the computational region, and it would also be of interest to show how 'well spread' the distribution is in R_q^*. Some relevant existing work that may prove useful for further investigation of this direction includes the work of Shparlinski and Banks [7] who used exponential sum tools to show that the inverses of small norm polynomials in R_q^* are well spread (although they could only prove this in the statistical region, i.e. for polynomials of norm $> q^{1/2+\varepsilon}$), and the work of Li and Roche-Newton, who use tools from additive combinatorics to show (see remark after Theorem 1.4 in [42]) a 'sum-product' type lower bound of the form $\Omega(|S|^{1+\delta})$ for some constant $\delta > 0$ on the size of the support of quotients of elements from a quite general class of subsets S of a finite field $\mathbb{F}$, in the 'computational region' where $|S| < |\mathbb{F}|^{1/2}$. Finally, we remark that besides allowing the use of smaller keys with $\sigma < q^{1/2}$ in the basic NTRU scheme and thus leading to improved efficiency, the hardness of the problem in the computational region also has several novel applications, as we shall see in Section 5.

The proof of [71] that the ratio g/f is close to uniform on R_q^* in the statistical region $\sigma > \text{poly}(n)q^{1/2}$, relies on sampling f and g from (modified) *discrete Gaussian* distributions χ_σ, and reducing the problem of showing the closeness to uniformity of h to showing the closeness to uniformity of a discrete gaussian distribution on $\mathbb{Z}^{2n}$ reduced modulo a certain NTRU-like lattice L. Using Fourier analysis, it was shown by Micciancio and Regev in [50] that the latter problem, for an arbitrary lattice L, can be reduced to upper bounding a geometric property of L called its *smoothing parameter*, which is closely related to the last Minkowski minimum of the lattice. The proof of [71] then reduces to bounding the smoothing parameter of the relevant NTRU-like lattices. Accordingly, we first introduce the discrete Gaussian and smoothing parameter tools from [50], that were further refined in [27, 56], and then come back to explain how they were applied in [71].

4.2 Gaussian Distributions Modulo Lattices and Fourier Analysis

For $x \in \mathbb{R}^n$, and $\sigma \in \mathbb{R}$, we let $\rho_\sigma(x) = e^{-\pi\|x\|^2/\sigma^2}$ denote a Gaussian function with deviation parameter σ evaluated at x, and $\nu_\sigma(x) = \sigma^{-n} \cdot \rho_\sigma(x)$ the density function of a continuous Gaussian random variable on $\mathbb{R}^n$ with parameter σ (note that $\sigma/\sqrt{2\pi}$ is the usual standard deviation parameter). Given a lattice $L \subseteq \mathbb{Z}^n$ with basis B, Micciancio and Regev [50] studied the density function

$$\nu'_\sigma(x) \stackrel{\text{def}}{=} (\nu_\sigma \bmod L)(x) = \sum_{v \in L} \nu_\sigma(x+v) = \sigma^{-n} \cdot \sum_{v \in L} \rho_\sigma(x+v) \qquad (4.1)$$

on the fundamental parallelepiped $P(B)$ of L, obtained by reducing a sample from ν_σ modulo $P(B)$. Intuitively, as the width parameter σ of ν_σ increases beyond the diameter of the parallelepiped $P(B)$, reducing the distribution ν_σ modulo $P(B)$ should 'fill' $P(B)$ and result in an approximately uniform density for ν'_σ. To make this intuition

precise, Micciancio and Regev applied Fourier analysis. Namely, using the right-hand side of (4.1) as the definition of ν'_σ, which is well defined for all $x \in \mathbb{R}^n$, one can extend the domain of ν'_σ from $P(B)$ to all of $\mathbb{R}^n$. Since L is closed under addition, it is clear that the extended function ν'_σ is periodic on L, i.e. $\nu'_\sigma(x+v) = \nu'(x)$ for any $x \in \mathbb{R}^n$ and $v \in L$, and thus naturally has a Fourier series representation of the form

$$\nu'_\sigma(x) = \det(\hat{L})\sigma^{-n} \cdot \sum_{w \in \hat{L}} \hat{\rho_\sigma}(w) e^{2\pi i \langle x,w \rangle} = \det(\hat{L}) \cdot \sum_{w \in \hat{L}} \rho_{1/\sigma}(w) e^{2\pi i \langle x,w \rangle}, \quad (4.2)$$

where the $\hat{\rho_\sigma}(w) = \sigma^n \rho_{1/\sigma}(w)$ is the Fourier transform of ρ_σ evaluated at the vector w in the *dual* lattice $\hat{L}$. The Fourier coefficient $\det(\hat{L}) \cdot \rho_{1/\sigma}(0) = \det(\hat{L}) = 1/\det(L)$ corresponding to the zero vector of $\hat{L}$ contributes the uniform (constant) term of ν'_σ's Fourier series (4.2), whereas all other Fourier coefficients $\det(\hat{L}) \cdot \rho_{1/\sigma}(w)$ corresponding to the non-zero vectors $w \in \hat{L} \setminus \{0\}$ contribute non-uniform terms in equation (4.2). A natural measure of the non-uniformity of ν'_σ is therefore the sum of the non-zero Fourier coefficients

$$S_\sigma(L) = \sum_{w \in \hat{L} \setminus \{0\}} \rho_{1/\sigma}(w). \quad (4.3)$$

Indeed, we have, for all $x \in P(B)$ that

$$\nu'_\sigma(x) = \sigma^{-n} \cdot \sum_{v \in L} \nu_\sigma(x+v) \in \frac{1}{\det L} \cdot [1 - S_\sigma(L), 1 + S_\sigma(L)], \quad (4.4)$$

and it follows that the statistical distance Δ of ν'_σ from the uniform density on $P(B)$ is at most $\frac{1}{2} S_\sigma(L)$. As σ increases, the width $1/\sigma$ where most of the mass of $\rho_{1/\sigma}$ is concentrated, decreases and eventually drops below the length $\lambda_1(\hat{L})$ of the shortest non-zero vector of $\hat{L}$. Therefore, for $\sigma > 1/\lambda_1(\hat{L})$, all the terms in the sum $S_\sigma(L)$ are in the small 'tail' of $\rho_{1/\sigma}$, so that $S_\sigma(L)$ (and hence also Δ) starts to rapidly decrease with σ, and can be made as small as any $\varepsilon > 0$ by choosing σ sufficiently large. Micciancio and Regev called this phenomena the *smoothing* of the distribution ν' and were led to the define the *ϵ-smoothing parameter* $\eta_\varepsilon(L)$ of a lattice L as the smallest value of σ such that $S_\sigma(L) \leq \varepsilon$. To bound the smoothing parameter $\eta_\varepsilon(L)$ in the 'smoothing region', where $\varepsilon = 2^{-n}$ is negligible, Micciancio and Regev applied a lattice sum tail bound for the Gaussian function due to Banascyk [6], who proved that for any lattice $L \in \mathbb{R}^n$ and $s > 0$, we have

$$\sum_{v \in L \setminus \sqrt{n}s \cdot \mathcal{B}} \rho_s(v) < 2^{-2n} \cdot \sum_{v \in L} \rho_s(v), \quad (4.5)$$

where $\mathcal{B}$ denotes the unit Euclidean ball in $\mathbb{R}^n$. Decomposing the sum on the right-hand side of equation (4.5) into its contributions from points inside and outside $\sqrt{n}s\mathcal{B}$ gives $\sum_{v \in L \setminus \sqrt{n}s \cdot \mathcal{B}} \rho_s(v) < 2^{-n} \sum_{v \in L \cap \sqrt{n}s \cdot \mathcal{B}} \rho_s(v)$. Taking $s = 1/\sigma$ with $\sigma > \sqrt{n}/\lambda_1(\hat{L})$, we have $\hat{L} \setminus \sqrt{n}/\sigma \cdot \mathcal{B} = \hat{L} \setminus \{0\}$ and $\hat{L} \cap (\sqrt{n}/\sigma \cdot \mathcal{B}) = \{0\}$, which

allowed Micciancio and Regev to conclude that $S_\sigma(L) = \rho_{1/\sigma}(\hat{L} \setminus \sqrt{n}/\sigma) \leq 2^{-n}$, i.e. that $\eta_{2^{-n}}(L) \leq \sqrt{n}/\lambda_1(\hat{L})$. The transference bound $\lambda_1(\hat{L}) \geq 1/\lambda_n(L)$ then gives $\eta_{2^{-n}}(L) \leq \sqrt{n}\lambda_n(L)$. A more refined analysis in [50] for general $\varepsilon > 0$ and further refinement by Peikert [56] in terms of the infinity norm minimum $\lambda_1^\infty(\hat{L})$ gave the following result.

Lemma 4.2 ([56, Le. 3.5],[50, Le. 3.3]). *For any full-rank lattice $L \subseteq \mathbb{R}^n$ and $\varepsilon \in (0,1)$, we have*

$$\eta_\varepsilon(L) \leq \sqrt{\frac{\ln(2n(1+1/\varepsilon))}{\pi}} \cdot \min\left(\lambda_n(L), 1/\lambda_1^\infty(\hat{L})\right).$$

Micciancio and Regev also studied *discrete* Gaussian distributions on lattices. For a lattice $L \subseteq \mathbb{R}^n$, a real $\sigma > 0$ and a point $c \in \mathbb{R}^n$, they defined the *discrete Gaussian distribution* of support L, deviation σ and center c by $D_{L,\sigma,c}(x) = \frac{\rho_\sigma(x-c)}{\rho_\sigma(L-c)}$, for any $x \in L$. They showed using the above Fourier approach that, in the 'smoothing' region where $\sigma \geq 2\eta_\varepsilon(L)$ and ε is small, the statistical properties of the discrete Gaussian distribution $D_{L,\sigma,c}$ are similar to those of the corresponding continuous Gaussian distribution $\rho_{\sigma,c}$. For example, they showed that the first and second moments of $D_{L,\sigma,c}(x) = \frac{\rho_\sigma(x-c)}{\rho_\sigma(L-c)}$ are bounded as

$$|E_{x \hookleftarrow D_{L,\sigma,c}}[\langle x-c, u\rangle]| \leq \frac{\varepsilon\sigma}{1-\varepsilon} \tag{4.6}$$

and

$$E_{x \hookleftarrow D_{L,\sigma,c}}[|\langle x-c, u\rangle|^2] \in \frac{\sigma^2}{2\pi} \cdot \left[1 - \frac{2\pi\varepsilon}{1-\varepsilon}, 1 + \frac{2\pi\varepsilon}{1-\varepsilon}\right] \tag{4.7}$$

for a unit vector $u \in \mathbb{R}^n$, compared to the corresponding values 0 and $\frac{\sigma^2}{2\pi}$ for $\rho_{\sigma,c}$, and that the norm of samples from $D_{L,\sigma,c}$ is bounded as

$$\Pr_{x \hookleftarrow D_{L,\sigma,c}}[\|x-c\| \geq \sigma\sqrt{n}] \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot 2^{-n}, \tag{4.8}$$

compared to a bound 2^{-n} on the probability of the same event for x sampled from the continuous distribution $\rho_{\sigma,c}$.

Importantly for many subsequent applications and the next section, Gentry, Peikert and Vaikuntanathan [27] showed that the above 'smoothing' phenomenon (4.4) for the distribution of a continuous Gaussian $\rho_{\sigma,c}$ reduced modulo a lattice L', also holds for a discrete Gaussian $D_{L,\sigma,c}$ on a lattice L, when it is reduced modulo a sublattice L' for which $\sigma \geq \eta_\varepsilon(L')$.

Lemma 4.3 ([27, Cor. 2.8]). *Let $L' \subseteq L \subseteq \mathbb{R}^n$ be lattices. For any $c \in \mathbb{R}^n$, $\varepsilon \in (0, 1/2)$ and $\sigma \geq \eta_\varepsilon(L')$, we have $\Delta(D_{L,\sigma,c} \bmod L'; U(L/L')) \leq 2\delta$ (here, $U(L/L')$ denotes the uniform distribution on the finite quotient group L/L').*

In [27], the above result was applied with $L = \mathbb{Z}^m$ and $L' = \{x \in \mathbb{Z}^m : A \cdot x = 0 \bmod q\}$ for some matrix $A \in \mathbb{Z}_q^{n \times m}$, in order to bound the distance to uniformity of the distribution of $A \cdot x \bmod q$ over the choice of x sampled from $D_{\mathbb{Z}^m,\sigma,c}$. They then bounded the smoothing parameter of L' with high probability for a uniformly distributed matrix A in $\mathbb{Z}_q^{m \times n}$.

Before we leave our brief review of this topic, we remark that cryptographic applications need also to efficiently sample from such discrete Gaussian distributions. Some applications, such as the NTRU key generation described next, require only discrete Gaussians with support $\mathbb{Z}^n$, which can be realized as n independent one-dimensional discrete Gaussian samples on $\mathbb{Z}$. Other applications, such as identity-based encryption (see, e.g.[27]), typically require discrete Gaussians $D_{L,\sigma}$ on non-orthogonal lattices L, for which the coordinates are not independent. Nevertheless, it was shown in [27] that given a basis $B = (b_1, \ldots, b_n)$ for an n-dimensional lattice L, one can still sample from such Gaussians efficiently when σ is slightly greater than $\|B\| = \max_i \|b_i\|$ by reducing it recursively to the one-dimensional case, via an algorithm that is essentially a randomized version of the Babai nearest plane algorithm [5]. Using an improved variant of the latter algorithm [15], gives the following result.

Lemma 4.4 ([15, Le. 2.2]). *There exists a polynomial-time algorithm that takes as input any basis $(b_i)_i$ of any lattice $L \subseteq \mathbb{Z}^n$ and $\sigma \geq \sqrt{\ln(2n+4)/\pi} \max_i \|b_i\|$, and returns a sample from the distribution $D_{L,\sigma}$.*

Efficient sampling algorithms for discrete Gaussians were given recently in [22, 21, 18, 65].

4.3 Statistical Hardness of the NTRU Decision Key Cracking Problem

We are now ready to state the result of [71] on the statistical hardness of the decision NTRU key cracking problem in the statistical region, and sketch the main steps in its proof.

Theorem 4.5 ([71, 72]). *Let $n \geq 8$ be a power of 2 such that $\phi = x^n + 1$ splits into n irreducible factors modulo prime $q \geq 5$ and let $R = \mathbb{Z}[x]/(\phi(x))$. Let $\chi_\sigma = D^*_{\mathbb{Z}^n,\sigma}$ be the distribution of $f \in R$ obtained by sampling f from $D_{\mathbb{Z}^n,\sigma}$ and rejecting (and resampling f) if $f \notin R_q^*$. Let $0 < \delta < 1/3$ be a constant and suppose that*

$$\sigma \geq n \cdot \sqrt{\ln(8nq)} \cdot q^{1/2+\delta}. \tag{4.9}$$

Then the Decision Key Cracking Problem $\mathtt{DNKC}_{n,q,\phi,\chi_\sigma}$ is statistically hard. More precisely, the NTRU key distribution $D_0 = \{h = g/f \in R_q : f, g \hookleftarrow D^*_{\mathbb{Z}^n,\sigma}\}$ is within statistical distance Δ from the uniform distribution $D_1 = U(R_q^*)$, with

$$\Delta \leq 2^{10n} q^{-\lfloor \delta \cdot n \rfloor}. \tag{4.10}$$

Before we sketch the proof of [71], we remark that in [72], the authors present a generalization of this result to the case when $\phi = x^n + 1$ splits into any number $k_q \in \{1, \ldots, n\}$ of irreducible factors mod q. In particular, for the case when $k_q = O(1)$, they show that the factor n in the lower bound on σ in Equation (4.9) can be reduced to approximately $\sqrt{n}$. We also remark that, to simplify the exposition below, the above version of the result in [71, 72] omits the restriction $f = 1 \bmod p$ on the distribution χ_σ of f. This restriction can be readily handled by a suitable modification of the argument (we refer to [71, 72] for details).

The proof of Theorem 4.5 proceeds as follows. The goal is to bound the statistical distance $\Delta = \frac{1}{2}\sum_{a \in R_q^*} |\Pr_{f,g}[g/f = a] - |R_q^*|^{-1}|$ by some small amount ξ. To do that, it is sufficient to show that for an overwhelming majority of the $a \in R_q^*$, the corresponding term in the sum Δ satisfies

$$|\Pr_{f,g}[g/f = a] - |R_q^*|^{-1}| < |R_q^*|^{-1} \cdot \xi. \tag{4.11}$$

In turn, since the event $g/f = a$ is equivalent to $fa - g = 0$ and also to $faa' - ga' = 0$ for any $a' \in R_q^*$, the termwise condition (4.11) is equivalent to showing that for the overwhelming majority of the pairs $(a_1, a_2) \in (R_q^*)^2$, the probability $P_{(a_1,a_2)}(0) \stackrel{\text{def}}{=} \Pr_{f,g}[fa_1 + ga_2 = 0]$ satisfies

$$|P_{(a_1,a_2)}(0) - |R_q^*|^{-1}| < |R_q^*|^{-1} \cdot \xi. \tag{4.12}$$

But, $P_{(a_1,a_2)}(0)$ is just the probability that $(f, g) \in L_{a_1/a_2}$, where

$$L_{a_1/a_2} \stackrel{\text{def}}{=} \{(f, g) \in R^2 : fa_1 + ga_2 = 0 \bmod q\}, \tag{4.13}$$

is in fact the NTRU module (or lattice in $\mathbb{Z}^{2n}$) corresponding to $a_1/a_2 \in R_q^*$. Hence, a sufficient condition for satisfying (4.12) is that for the majority of (a_1, a_2), the distribution D of (f, g) reduced modulo the lattice L_{a_1/a_2}, is close (within statistical distance $\leq |R_q^*|^{-1} \cdot \xi$) to uniform on $\mathbb{Z}^{2n}/L_{a_1/a_2}$, over the choice of (f, g) sampled from $(D^*_{\mathbb{Z}^n,\sigma})^2$. Now, the latter problem can almost directly be attacked using the 'smoothing modulo a lattice' Lemma 4.3, with $L = \mathbb{Z}^{2n}$ and $L' = L_{a_1/a_2}$, reducing the problem to that of bounding the smoothing parameter of the NTRU lattice L_{a_1/a_2} for the majority of (a_1, a_2). Indeed, the only issue preventing this direct application of Lemma 4.3 here, is that the distribution $(D^*_{\mathbb{Z}^n,\sigma})^2$ of (f, g) is supported on a set $\mathbb{Z}^{2n} \cap (R_q^*)^2$ which, unlike $\mathbb{Z}^{2n}$, is *not* a lattice. Indeed, by the choice of q, we have $\phi(x) = \prod_{i=1}^n \phi_i(x)$ where $\phi_1, \ldots, \phi_n$ denote the linear factors of ϕ mod q, and thus the Chinese Remainder Theorem gives the isomorphisms $R_q \simeq \prod_{i=1}^n \mathbb{Z}_q[x]/(\phi_i(x)) \simeq \mathbb{F}_q^n$ and $R_q^* \simeq \prod_{i=1}^n (\mathbb{Z}_q[x]/(\phi_i(x)) \setminus \{(q, \phi_i(x))\}) \simeq (\mathbb{F}_q^*)^n$. Accordingly, Stehlé and Steinfeld handled this issue by decomposing $\mathbb{Z}^{2n} \cap (R_q^*)^2$ in terms of lattices as follows:

$$\mathbb{Z}^{2n} \cap (R_q^*)^2 = \left(\mathbb{Z}^n - \bigcup_{S \subseteq \{1,\ldots,n\}} I_S \right)^2, \tag{4.14}$$

where for a subset $S \subseteq \{1, \ldots, n\}$, I_S denotes the ideal of R_q generated by $\prod_{i \in S} \phi_i(x)$. Similarly, they decomposed the non-lattice set $L^*_{a_1/a_2} \stackrel{\text{def}}{=} L_{a_1/a_2} \cap (R^*_q)^2$ in terms of lattices as:

$$L^*_{a_1/a_2} \cap (R^*_q)^2 = L_{a_1/a_2} - \bigcup_{S \subseteq \{1,\ldots,n\}} L_{a_1/a_2}(I_S), \tag{4.15}$$

where $L_{a_1/a_2}(I_S)$ denotes the sublattice of L_{a_1/a_2} consisting of pairs (f, g) in which both f and g belong to the same ideal I_S of R_q (note that if S is the empty set, we have $L_{a_1/a_2}(I_S) = L_{a_1/a_2}$). Thus, the numerator and denominator in the probability

$$P_{(a_1,a_2)}(0) = D_{\mathbb{Z}^{2n},\sigma}(L^*_{a_1/a_2}) / D_{\mathbb{Z}^{2n},\sigma}((R^*_q)^2), \tag{4.16}$$

could be decomposed by applying the inclusion-exclusion principle to the set decompositions (4.15) and (4.14), giving respectively:

$$D_{\mathbb{Z}^{2n},\sigma}(L^*_{a_1/a_2}) = \sum_{S \subseteq \{1,\ldots,n\}} (-1)^{|S|} \cdot D_{\mathbb{Z}^{2n},\sigma}(L_{a_1/a_2}(I_S)), \tag{4.17}$$

and

$$D_{\mathbb{Z}^{2n},\sigma}((R^*_q)^2) = \left(\sum_{S \subseteq \{1,\ldots,n\}} (-1)^{|S|} \cdot D_{\mathbb{Z}^n,\sigma}(I_S) \right)^2. \tag{4.18}$$

Each term $D_{\mathbb{Z}^{2n},\sigma}(L_{a_1/a_2}(I_S))$ (resp. $D_{\mathbb{Z}^n,\sigma}(I_S)$) in the sum (4.17) (resp. (4.18) is now in the form of a discrete Gaussian evaluated on a a lattice. This allowed Stehlé and Steinfeld to apply smoothing Lemma 4.3 to each term for which the smoothing condition $\eta_\varepsilon(L_{a_1/a_2}(I_S)) \leq \sigma$ holds and conclude that $D_{\mathbb{Z}^{2n},\sigma}(L_{a_1/a_2}(I_S))$ is close to $1/\det(L_{a_1/a_2}(I_S)) = q^{-(n+|S|)}$ for those terms. For other terms with $|S|$ sufficiently large so that $\eta_\varepsilon(L_{a_1/a_2}(I_S))$ larger than σ and the smoothing condition does not hold, they used the fact that $L_{a_1/a_2}(I_S)$ is a sub-lattice of $L_{a_1/a_2}(I_{S'})$ for some $S' \subseteq S$ with $|S'| \geq \varepsilon n$ for which the smoothing condition does hold, so that the bound $D_{\mathbb{Z}^{2n},\sigma}(L_{a_1/a_2}(I_S)) \leq D_{\mathbb{Z}^{2n},\sigma}(L_{a_1/a_2}(I_{S'})) \approx 1/\det(L_{a_1/a_2}(I_S))$ holds, with the latter approximation obtained by the smoothing Lemma applied to $L_{a_1/a_2}(I_{S'})$. Overall, assuming the smoothing condition

$$\sigma \geq \eta_\varepsilon(L_{a_1/a_2}(I_S)) \text{ for } |S| \leq \varepsilon' n \text{ and } \varepsilon \leq \delta q^{-n} \tag{4.19}$$

holds for $|S| \leq \varepsilon' n$ with some constant $\varepsilon' > 0$, this allowed approximating the sum (4.17) as

$$D_{\mathbb{Z}^{2n},\sigma}(L^*_{a_1/a_2}) \approx \sum_{k=0}^{n} (-1)^k q^{-(n+k)} = \left(\sum_{k=0}^{n} (-1)^k q^{n-k}\right) q^{-2n} = (q-1)^n q^{-2n}, \tag{4.20}$$

with an approximation error bounded as $2^{O(n)}q^{1-(1+\varepsilon)n}$. A similar argument was applied to (4.18) to obtain the denominator approximation

$$D_{\mathbb{Z}^{2n},\sigma}((R_q^*)^2) \approx |R_q^*|^2 q^{-2n}, \tag{4.21}$$

so that $|P_{(a_1,a_2)}(0) - 1/|R_q^*|| \leq 2^{O(n)}q^{1-(1+\varepsilon')n}$, which leads to a statistical distance bound of the desired form $2^{O(n)}q^{-\varepsilon' n}$.

The remaining step in the proof of Theorem 4.5 was to show that $\sigma \geq n \cdot \sqrt{\ln(8nq)} \cdot q^{1/2+\delta}$ for some small $\delta > 0$ is sufficient to satisfy the smoothing condition (4.19). To do so, Stehlé and Steinfeld proved an upper bound on the smoothing parameter $\eta_\varepsilon(L_{a_1/a_2}(I_S))$ of the NTRU-like lattices $L_{a_1/a_2}(I_S)$, that holds with overwhelming probability over random a_1, a_2 in R_q^*.

Lemma 4.6 ([72, Le.3.2 and Le 2.1]). *Let $n \geq 8$ be a power of 2 such that $\phi = x^n + 1$ splits into n linear factors modulo a prime $q \geq 5$ and let $S \subseteq [n]$ with $|S| \leq \varepsilon' n$. Then for all except a fraction $\leq 2^{4mn}q^{-\varepsilon mn}$ of $a \in (R_q^*)^m$, we have $\eta_\varepsilon(L_{a_1/a_2}(I_S)) \leq \sqrt{n\ln(2mn(1+1/\varepsilon))/\pi} \cdot q^{\frac{1}{2}+\varepsilon'/2+\varepsilon}$.*

Taking $\varepsilon \leq \xi q^{-n}$ for some small ξ in this Lemma gives $\eta_\varepsilon(L_{a_1/a_2}(I_S)) \leq n \cdot \sqrt{\ln(8nq)} \cdot q^{1/2+\delta}$ for a small δ, as required. Lemma 4.6 was proved by first applying Lemma 4.2 to reduce the problem of bounding the smoothing parameter of the generalized NTRU lattice $L_{a_1/a_2}(I_S)$ to the problem of lower bounding the minimum of the *dual lattice* $\widehat{L_{a_1/a_2}}(I_S)$. As shown in [71], this dual has a simple description of the form $\widehat{L_{a_1/a_2}}(I_S) = \frac{1}{q} \cdot \left\{(t_1,t_2) \in R^2 : \exists s \in R_q, \forall i, (t_i \bmod q) = a_i \cdot s \bmod I_S\right\}$. They then applied a counting argument to lower bound the minimum of this lattice, upper bounding the number of (a_1, a_2) for which $\widehat{L_{a_1/a_2}}(I_S)$ contains a given fixed short vector and applying the union bound over all such 'bad' pairs (a_1, a_2). Due to the simple description of $\widehat{L_{a_1/a_2}}(I_S)$ and the CRT isomorphism $R_q^* \simeq (\mathbb{F}_q^*)^n$, the latter counting problem reduced to the easy problem of counting the number of solutions to linear equations over the field $\mathbb{F}_q$.

4.4 Computational Hardness of the Ciphertext Cracking Problem

We now turn to explain the main ingredients in second part of the proof of Theorem 4.1, i.e. the computational reduction between the decisional NTRU ciphertext cracking problem $\texttt{DNCC}_{n,q,\phi,\chi_\rho,\chi_\beta}$ and the worst-case hardness of γ-Ideal-SVP$_\phi$. This reduction is essentially a special case of the hardness proof for a problem now known in the field of lattice-based cryptography as *Ring Learning With Errors* (Ring-LWE) problem. The decision and search variants of quite general variants of the problem were shown hard by Lyubashevsky et al. [48], while a concurrent and independent work by Stehlé et al. [73] showed an alternative (and conceptually more modular)

proof of hardness for a search variant. We remark that since its introduction by Regev in 2005 [61], the LWE problem and its variants have found many interesting cryptographic applications; we refer the reader to [51] for a partial survey of applications, and to the survey by Regev [62] on the hardness of the Ring-LWE problem and its more general Learning with Errors (LWE) relative. Here, we only briefly summarize the connection to and results relevant to NTRU.

Ring-LWE Problem. Let R_q be as above. The *decision Ring-LWE problem*, denoted by R-LWE$_{n,q,\phi,\chi_\beta,r}$, for R_q and 'small' noise distribution χ_β is to distinguish, given $r = \text{poly}(n)$ independent identically distributed samples $(a_i, b_i) \in R_q^2$ for $i = 1, \ldots, r$, whether these samples were sampled from the distribution $D_0(s)$ (defined in the following) or the uniform distribution $U(R_q^2)$ on pairs of elements from R_q. The distribution $D_0(s)$ is defined as follows. For some uniformly distributed 'secret' $s \in R_q$ (chosen once and for all, i.e. the same s is used for all r samples from $D_0(s)$), we sample (a_i, b_i) from distribution $D_0(s)$ by sampling a_i uniformly from R_q and noise term e_i from the distribution χ_β concentrated on 'small' elements, and setting $b_i = a_i \cdot s + e_i \in R_q$. The *search Ring-LWE problem* is the defined in a similar way, except that the given samples are always sampled from the distribution $D_0(s)$, and the goal is to compute s (and hence also the e_i's).

The following hardness result for the decision Ring-LWE problem is adapted from the results in [48]. It works with the noise distribution D_β on R defined by sampling a continuous e'_i from the continuous Gaussian distribution ν_β with deviation parameter β and rounding it to R.

Theorem 4.7 (Adapted from [48]). *Fix $\varepsilon > 0$, and let n be a power of 2 such that $\phi = x^n + 1$ splits into n irreducible factors modulo prime $q = \text{poly}(n)$, let $\chi_\beta = D_\beta$ be the rounded Gaussian distribution with parameter $\beta = n^{1.25+\varepsilon} < q$, and assume that $r = O(1)$. Then there exists a randomized polynomial-time quantum reduction from γ-*Ideal-SVP *to* R-LWE$_{n,q,\phi,\chi_\beta,r}$, with $\gamma = O(n^{1.75+\varepsilon}) \cdot \frac{q}{\beta}$.*

In the NTRU application described below, the above hardness result is applied with $r = 2$, which is the smallest number of samples for which the Ring-LWE problem may be statistically solved. With somewhat more complex Gaussian-like noise distributions χ_β, it is shown in [48] that one can handle a larger number $r = \text{poly}(n)$ of samples and obtain slightly better reduction parameters (such a distribution is used in [71, 72]).

Relation to Ciphertext Cracking Problems. We can now explain the connection observed in [71, 72] between the NTRU ciphertext cracking problem and the Ring-LWE problem. The `DNCC` distinguisher algorithm A can distinguish, given h uniform in R_q^*, whether c comes from the distribution D_0 with $c = p \cdot (hs + e)$, s sampled from χ_ρ and e sampled from χ_β, or c comes from $D_1 = U(R_q)$. Since $p \in R_q^*$, we can modify A to A' that can distinguish D'_0 from D_1, where in D'_0, $c = hs + e$ with s sampled from χ_ρ and e sampled from χ_β (indeed, given c sampled from either D'_0 or D_1, the modified distinguisher A' maps c to $c' = p \cdot c \in R_q$, and runs the original distinguisher A on input c' – this maps distribution D_0 to D'_0 and D_1 to itself). This

latter ciphertext cracking problem is of the same form as a single Ring-LWE sample, except that here s comes from a 'small' distribution χ_β instead of being uniform in R_q as in the Ring-LWE problem.

However, as observed in [4], the standard Ring-LWE problem with $r = 2$ samples and noise distribution χ_β (but s uniform in R_q) can be easily reduced to the ciphertext cracking Ring-LWE variant with one sample but small s, sampled independently from the same distribution χ_β as the noise e (this variant has been termed by [4] the 'Hermite Normal Form' variant of LWE). Indeed, one can map an instance of the former $(a_1, b_1), (a_2, b_2)$ to $(h = -a_1^{-1} \cdot a_2, c = b_2 - a_1^{-1} b_1 a_2)$ (note that h is uniform in R_q^* when a_1, a_2 are, and $c = he_1 + e_2$ when b_1, b_2 come from D_0, whereas c is uniform in R_q and independent of h when b_1, b_2 come from $U(R_q)$). The only remaining difference between the latter Ring-LWE variant with $r = 2$ and the standard Ring-LWE problem is that here, the distribution of the a_i is $U(R_q^*)$ instead of $U(R_q)$. But, since the probability that (a_1, a_2) sampled from $U(R_q^2)$ falls in $(R_q^*)^2$ is $(1 - 1/q)^{2n} \geq 1 - 2n/q$ and this is non-negligible for $q > 2n$, the hardness of the standard Ring-LWE variant implies the hardness of the new variant. Stehlé and Steinfeld were therefore able to obtain the following relation.

Lemma 4.8 (Adapted from [71]). *Let n be a power of 2 such that $\phi = x^n + 1$ splits into n irreducible factors modulo prime $q = \text{poly}(n)$, $p \in R_q^*$ and $\chi_\rho = \chi_\beta$. Then there exists a randomized polynomial-time reduction between the decision NTRU ciphertext cracking problem $\texttt{DNCC}_{n,p,q,\phi,\chi_\rho,\chi_\beta}$ and the decision Ring-LWE problem* $\text{R-LWE}_{n,q,\phi,\chi_\beta,r}$ *with $r = 2$ samples.*

Combining Lemma 4.8 and Theorem 4.7 allowed Stehlé and Steinfeld to obtain the hardness of the decision key cracking problem based on worst-case hardness of poly(n)-Ideal-SVP.

5 Recent Developments in Applications of NTRU

In this Section, we review two recent novel applications of the NTRU system, that add extra powerful functionality to the basic NTRU encryption scheme, and provide fresh motivation for the study of both old and new variants of the NTRU problem.

5.1 NTRU-Based Homomorphic Encryption

A *homomorphic* encryption scheme allows any party Bob holding ciphertexts $c_1 = \mathsf{Enc}_{pk_A}(m_1), \ldots, c_t = \mathsf{Enc}_{pk_A}(m_t)$ of some messages $m_1, \ldots, m_t$ encrypted with Alice's public key pk_A, to compute a ciphertext $c = \mathsf{Enc}_{pk_A}(m)$ for a message $m = f(m_1, \ldots, m_t)$ that is some function f of the messages $m_1, \ldots, m_t$. Here, the function f can be chosen by and known to Bob, whereas Bob may know nothing about the input messages $m_1, \ldots, m_t$ or output message m that he is processing. Only Alice, holding the secret key sk_A can decrypt the ciphertexts to access the messages.

Given a sufficiently large class of allowed functions f, homomorphic encryption can enable a range of exciting privacy-enhanced applications, including secure outsourced computations for 'cloud-based' computing, private database queries, and others. The concept of homomorphic encryption schemes was proposed in the 1970's by Rivest et al. [64], but until recently, all proposed realizations were very limited in the class of of functions f allowed.

In a major breakthrough in 2009, Gentry [25] proposed the first plausible candidate for a *fully* homomorphic encryption (FHE) scheme, allowing f to be an arbitrary function. Gentry's scheme was based on problems on ideal lattices, with the scheme being naturally homomorphic with respect to the underlying ring operations. Gentry's original scheme was quite inefficient, and several improved schemes based on the LWE and Ring-LWE problem have now been proposed (e.g. [16, 14, 13, 28]), but one of the simplest and possibly more practical candidates is the NTRU based fully homomorphic encryption scheme proposed by Lopez-Alt et al. [43]. The scheme proposed in [43] even has the novel feature that the ciphertexts combined homomorphically could have been encrypted to multiple recipients ('multikey fully homomorphic encryption'), so that the resulting ciphertext c can be decrypted only jointly by all of these recipients together. This multikey feature has potentially interesting applications in secure multiparty computation protocols, but for simplicity, we focus below on the single key variant of this scheme and refer the interested reader to [43] for details on the multikey variant.

The starting point of the NTRU-based FHE scheme of [43] is the observation that the basic NTRU encryption scheme has natural homomorphic properties with respect to both addition and multiplication in the ciphertext ring R_q, corresponding, respectively, to addition and multiplication in the plaintext ring R_p (in the following, one can think of the case $p = 2$; in this case R_p contains a subring isomorphic to the binary field $\mathbb{F}_2$. Since any arbitrary function f can be written as a circuit over $\mathbb{F}_2$, a scheme homomorphic over R_p is fully homomorphic). That is, given two NTRU ciphertexts $c_1 = h \cdot s_1 + pe_1 + m_1 \in R_q$ and $c_2 = h \cdot s_2 + pe_2 + m_2 \in R_q$ for messages $m_1, m_2 \in S_p$ with respect to public keys $h = pg/f \in R_q^*$, we have

$$c_1 + c_2 = h \cdot (s_1 + s_2) + p(e_1 + e_2) + (m_1 + m_2)$$

is an NTRU ciphertext with $s = s_1 + s_2$, $e = e_1 + e_2$ that decrypts to message $f \cdot (c_1 + c_2) \bmod p = m_1 + m_2 \bmod p$ if the decryption condition $\|p(gs + fe) + m\|_\infty < q/2$ holds. Similarly, we also have

$$c_1 \cdot c_2 = h^2 s_1 s_2 + h(s_1 e_2' + s_2 e_1') + p(e_1 e_2 + e_1 m_2 + e_2 m_1) + m_1 m_2, \qquad (5.1)$$

where $e_i' = pe_i + m_i$ for $i = 1, 2$. This ciphertext can in fact be decrypted with the secret key f^2 to give the message $f^2 \cdot (c_1 \cdot c_2) \bmod p = m_1 \cdot m_2 \bmod p$ as long as the decryption condition

$$\|(pg)^2 s_1 s_2 + (pfg)(s_1 e_2' + s_2 e_1') + f^2(p(e_1 e_2 + e_1 m_2 + e_2 m_1) + m_1 m_2)\|_\infty < q/2$$

holds. Note that the latter decryption condition is the bottleneck, and that if $\|pg\|_\infty \approx \|f\|_\infty \approx \sigma$ and $\|s_i\|_\infty \approx \|e_i\|_\infty \beta$ then the decryption condition is of the form $(\sigma \cdot \beta \text{poly}(n))^2 \leq q/2$, compared to the condition $(\sigma \cdot \beta \text{poly}(n) \leq q/2$ for the basic (not-homomorphic) NTRU scheme, with the multiplicative homomorphism contributing the squared term. As a consequence, even to be homomorphic for just one multiplication, the NTRU scheme requires $\sigma\beta < q^{1/2}/\text{poly}(n)$, which necessitates the hardness of the key cracking problem in the computational region, where $\sigma < q^{1/2}/\text{poly}(n)$, and gives new motivation for studying the hardness of this problem.

The above homomorphic property naturally extends to more than one multiplication, where each additional multiplication approximately squares the norm of the noise terms in decryption (ignoring the $\text{poly}(n)$ terms). As a consequence, the scheme can be made homomorphic with respect to polynomial functions of multiplicative *depth* up to d if $(\sigma\beta\text{poly}(n))^{2^d} < q/2$. With $q = 2^{n^\varepsilon}$ and $\sigma = \beta = \text{poly}(n)$, this allows the depth d to be up to $O(\log n)$, but with an even larger 'computational region' ratio $q^{1/2}/\sigma = \Omega(2^{n^\varepsilon}/\text{poly}(n))$, which weakens the security of the key cracking problem against approximate shortest vector problem algorithms (which become easy to solve in $\text{poly}(n)$ time when $\varepsilon \geq 1$). To allow larger homomorphic multiplicative depth d, the authors of [43] adapt techniques originally developed in [14] for Ring-LWE based FHE schemes, called *relinearization* and *modulus reduction*, which we now briefly sketch.

The first obstacle in the above basic homomorphic NTRU scheme, is the exponential growth in the degree 2^d of the secret key f^{2^d} needed to decrypt a ciphertext after d homomorphic multiplications, which implies a doubly-exponential growth in the secret key norm $\|f^{2^d}\|$. To avoid this growth, the authors modified the scheme to use the following relinearization procedure. The relinearization procedure is applied to a ciphertext $c = c_1 \cdot c_2$ after each homomorphic multiplication operation as in (5.1), to produce a new ciphertext $\hat{c}$ that encrypts the same message $m_1 \cdot m_2 \bmod p$ as c, but $\hat{c}$ can be decrypted with the original secret key f, rather than with the squared secret f^2 which is needed to decrypt c. To achieve this, the scheme must be modified in the key generation stage; in addition to publishing her NTRU public key $h = pg/f \in R_q^*$, the recipient Alice also publishes $\approx \log q$ additional ring elements ζ_τ that consist of 'pseudo-encryptions' of her squared secret key f^2:

$$\zeta_\tau = h \cdot s_\tau + pe_\tau + 2^\tau f^2 \in R_q \text{ for } \tau = 0, \ldots, \lfloor \log q \rfloor, \tag{5.2}$$

where for each τ, s_τ and e_τ are sampled from χ_β as in the usual NTRU encrption algorithm. Note that ζ_τ has the form of an encryption of $2^\tau f^2$, but it is not quite a valid encryption of this value since 2^τ is 'large' and thus typically outside the message space S_p (hence the name 'pseudo-encryption'). Nevertheless, the extra public ζ_τ allow the product ciphertext $c = c_1 \cdot c_2$ of the form (5.1) to be relinearized as follows. Let $c = \sum_\tau c_\tau 2^\tau$ denote the binary representation of c, where for each $\tau \in \{0, \ldots, \lfloor \log q \rfloor\}$, $c_\tau \in R_q$ is a $\{0, 1\}$ coefficient polynomial whose coefficients consist of the τ'th bit

in the binary representation of the coefficients of c. The relinearization procedure computes the new ciphertext

$$\hat{c} = \sum_{\tau} c_\tau \cdot \zeta_\tau = h \cdot (\sum_{\tau} c_\tau s_\tau) + p \cdot (\sum_{\tau} c_\tau e_\tau) + f^2 \cdot (\sum_{\tau} c_\tau 2^\tau). \tag{5.3}$$

Now, recall that the term $f^2 \cdot (\sum_\tau c_\tau 2^\tau) = f^2 c$ has the decryption form $pe + m_1 m_2$ for some small e, so in fact $\hat{c}$ has the form $\hat{c} = h \cdot \hat{s} + p \cdot \hat{e} + m_1 \cdot m_2$ of a standard NTRU ciphertext for $m_1 \cdot m_2$ mod p decryptable with f (rather than f^2) as required, where $s = \sum_\tau c_\tau s_\tau$ and $\hat{e} = \sum_\tau c_\tau e_\tau + e$ are small thanks to the smallness of the binary coefficients of c_τ.

The second obstacle that remains even in the above relinearized scheme, is that the norm $O(\text{poly}(n)^{2^d})$ of the ciphertext terms $\hat{s}$ and $\hat{e}$ in the relinearized ciphertext $\hat{c} = h \cdot \hat{s} + p \cdot \hat{e} + m_1 \cdot m_2 \in R_q$ still grows doubly exponentially in the multiplicative depth d. The *modulus reduction* technique used in [43] addresses this problem by applying an additional transformation to the relinearized ciphertext $\hat{c}$ after each homomorphic multiplication. Namely, modulus reduction scales down the ciphertext $\hat{c}$ over R_q to another ciphertext $\hat{c}' \in R_{q'}$ over a ring $R_{q'}$ with a smaller modulus $q' < q$, such that the noise terms in $\hat{c}'$ are also approximately scaled down by the ratio q'/q, while preserving the secret key $f \in R$ and the encrypted message $m_1 \cdot m_2$ mod p. By choosing $q'/q = 1/\text{poly}(n)$, the scaling down ratio q'/q can compensate for the $\text{poly}(n)$ growth ratio due to the homomorphic multiplication, so that overall, the norm of the noise terms does not grow after a homomorphic multiplication. The 'catch', of course, is that each multiplication reduces the modulus q of the underlying ciphertext space R_q by a $\text{poly}(n)$ factor, so that the modulus decreases exponentially as $q_0/\text{poly}(n)^d$ with the depth d, where q_0 denotes the initial modulus, and thus the process can only work for $d = O(\log q / \log n)$. Nevertheless, this exponential scaling of the modulus with the depth d is a big improvement over the doubly-exponential noise norm growth with d in the basic scheme, allowing $d = O(n^\varepsilon / \log n)$ for $q = 2^{n^\varepsilon}$, instead of $d = O(\log n)$ for the basic scheme. Moreover, using a *bootstrapping* technique originally due to Gentry [25], it is shown in [43] that the depth d achievable with this improved scheme can be leveraged to realize *fully* homomorphic encryption, i.e. homomorphic computation for functions of unbounded multiplicative depth.

Before we leave this topic, we point out that, besides its reliance on the hardness of the NTRU key cracking problem in the computational region (as already observed above), the security of the homomorphic NTRU scheme employing the relinearization technique above in fact relies on a stronger new *circular security* variant of the problem, that may be easier than the classical cracking problems.

Decision Circular Key Cracking Problem $\texttt{DNCKC}_{n,p,q,\phi,\chi_\sigma,\chi_\beta,\ell}$: Given (n, p, q, ϕ) and $(h, \{\zeta_\tau\}_\tau)$, distinguish whether $(h, \{\zeta_\tau\}_{\tau \le \ell})$ is sampled from the distribution $D_0 = \{(h = g/f \in R_q, \zeta_\tau = h \cdot s_\tau + pe_\tau + 2^\tau f^2 \in R_q : f, g \hookleftarrow \chi_\sigma, s_\tau, e_\tau \hookleftarrow \chi_\beta, 1 \le \tau \le \ell)\}$ or from the uniform distribution $D_1 = U(R_q^*) \times U(R_q^\ell)$.

Relating the hardness of this problem to the standard decision key cracking problem,

or giving an efficient algorithm for this problem, are interesting open problems. The term 'circular security' comes from the study of the security of encryption schemes that encrypt functions of their own secret key. Some results are known in this area for lattice based encryption schemes [4], but they do not seem directly applicable to the above NTRU variant of this problem.

Finally, with respect to assumptions required for realizing FHE, we remark that, by adapting techniques introduced by Brakerski for LWE and Ring-LWE based FHE schemes, it was recently shown by Bos et al. [12] how to modify the NTRU-based FHE scheme above to avoid the need for the hardness of the NTRU key cracking problem in the computational region, allowing the use of keys generated in the statistical region, for which the results of Section 4.3 could be applied (a more efficient variant that does need the computational region assumption is also presented in [12]). In terms of the lattice problem approximation factor needed for security, it was recently shown in [17] that one can construct FHE schemes based on worst-case lattice problems with a polynomial approximation factor asymptotically approaching (within an arbitrarily small polynomial factor) that of known (non-FHE) public-key encryption schemes.

5.2 NTRU-Based Multilinear Maps

At around the year 2000, the new powerful tool of bilinear maps (also known as pairings) was introduced into the field of public-key cryptography, and soon found many interesting applications, including non-interactive key agreement protocols [67, 38], identity-based encryption [9], and many others. While the original realization of bilinear maps was based on algebraic curves, researchers soon began to search for other realizations, and explored the fascinating cryptographic consequences of generalizations to multilinear maps [10]. Until very recently, however, candidate realizations of such multilinear remained elusive. But in a breakthrough result announced in 2012, Garg, Gentry and Halevi [23] showed that a functionality essentially equivalent to (and to some extent even more powerful than) the sought-after multilinear maps, can be achieved using a suitable variant of the NTRU encryption scheme. However, similarly to the NTRU-based FHE scheme from the previous section, the security of their so-called *graded encoding system* relies on the hardness of new and not yet well understood variants of the NTRU problems. In the following, we sketch a simplified variant of the system proposed in [23], which is due to Langlois et al. [40]. We believe that the simplified variant is more closely related to the NTRU scheme than the original version in [23].

We first informally review the main requirements of a k-graded encoding scheme over a ring R_p. Given some public parameters pk, the scheme has an efficient randomized *sampling* algorithm Samp that outputs a representative of an (almost) uniformly distributed 'level 0' element $m \in R_p$. There is also a (possibly randomized) *encoding* algorithm Enc that takes a 'level 0' element $m \in R_p$ and outputs a 'level 1 encoding' $c_1 = \mathsf{Enc}_1(m)$ of m. One can think of an encoding $\mathsf{Enc}_1(m)$ of m as similar

to a homomorphic encryption of m, in the sense that it should be hard to recover m from its encoding, and the encoding algorithm should have additive and multiplicative homomorphic properties up to a multiplicative depth ('level') k, i.e. there exist efficient algorithms add and mul with the following properties. Given a level i encoding $c_1 = \mathsf{Enc}_i(m)$ and level j encoding $c_2 = \mathsf{Enc}_j(m_2)$ for level 0 elements $m_1, m_2 \in R_p$, we have that $\mathsf{add}(\mathsf{par}, c_1, c_2) = \mathsf{Enc}_i(m_1 + m_2)$ is a level i encoding of $m_1 + m_2 \in R_p$ (here, we assume that $j = i$), while $\mathsf{mul}(\mathsf{par}, c_1, c_2) = \mathsf{Enc}_{i+j}(m_1 + m_2)$ is a level $i+j$ encoding of $m_1 \cdot m_2 \in R_p$. However, there is one major difference between graded encodings and homomorphic encryption schemes: unlike an encryption of element m, for correct functionality similar to that provided by multilinear maps, the encoding $\mathsf{Enc}_i(m)$ for $i \leq k$ should *not* satisfy semantic security, i.e. it should not hide all partial information on m (on the other hand, at level $\geq k+1$ the encodings *should* hide information on the encoded element; see the k-graded decision Diffie-Hellman problem below). In particular, given an encoding $c = \mathsf{Enc}_1(m)$ and the encoded element m, it should be easy to verify that c is indeed an encoding of m, rather than of some other element m'. In fact, there should exist an efficient 'extraction' algorithm Ext, such that given a level k encoding $c = \mathsf{Enc}_k(m)$ of m and public parameters pk, outputs a 'canonical and random' representative $r(m) = \mathsf{Ext}(pk, c) \in \{0,1\}^\ell$ of m, where ℓ should be proportional to the security parameter for the scheme. Namely, the extracted representative $r(m)$ should be dependent only on m and not on the randomness in the encoding of m, and for a uniformly distributed element $m \in R_p$, the extracted representative should be almost uniformly distributed in $\{0,1\}^\ell$.

A classical example application [23] of such a k-graded encoding scheme is constructing a non-interactive $(k+1)$-party key agreement protocol, generalizing the Diffie-Hellman 2-party protocol [20] and the 3-party protocol of Joux [38]. In this case, for $i = 1, \ldots, k+1$, party i privately chooses an element $m_i \in R_p$ and broadcasts the level 1 encoding $c_i = \mathsf{Enc}_1(m_i)$. The agreed shared key is $K = \mathsf{Ext}(pk, \mathsf{Enc}_k(m_1 m_2 \cdots m_{k+1})$, is the canonical representative of the element $m_1 \cdots m_{k+1}$. The key K can be computed by the ith party, by first multiplying the level 1 encodings of all other parties to get a level k encoding c' of $\prod_{j \neq i} m_j$, and then using its private level 0 element m_i to compute a level k encoding $c = m_i \cdot c$ of $m_1 \ldots m_{k+1}$, followed by applying Ext. An eavesdropping adversary, on the other hand, has to solve the *k-graded decision Diffie-Hellman problem*: given level 1 encodings $c_i = \mathsf{Enc}_1(m_i)$ for $i = 1, \ldots, k+1$ for uniformly distributed $m_i \in R_p$, distinguish $r(m_1 \cdots m_{k+1})$ from a random string. One could hope that this problem is as hard as the k-graded Discrete Log problem: given $c = \mathsf{Enc}_1(m)$ for a uniformly distributed $m \in R_p$, find m.

The NTRU-based construction of [23] for a k-graded encoding scheme, as simplified by [40], works as follows. The parameter generation algorithm is similar to the one for the basic NTRUEncrypt scheme, but with the following modifications. The public key is still of the form $h_i = pg_i/f \in R_q$ with 'small' $g_i, f \in R$ being sampled from a distribution χ_σ and f subject to the restriction $f = 1 \bmod p$. However, there are m_r such keys published, sharing the same denominator f but having independent

g_i for $i = 1, \ldots, m_r$[1]. Moreover, the choice of the polynomial p defining the encoded element ring R_p is different: instead of being a small public integer, p is chosen as a *secret* small polynomial from some high entropy distribution χ_p. To facilitate the extraction algorithm Ext, an additional element e_k is published, where

$$e_k = uf^k/p \in R_q,$$

for some u of norm $\|u\| = \text{poly}(n) \cdot q^{1/2}$. The 'level 0' sampling algorithm Samp samples $m \in R$ from a discrete Gaussian $D_{\mathbb{Z}^n,s}$ with s chosen small compared with q but sufficiently larger than the smoothing parameter $\eta_\varepsilon((p)) = O(\text{poly}(n) \cdot \|p\|)$ of the ideal of R generated by p, so that by the smoothing Lemma 4.3, the level 0 element m mod p is close to uniform on R_p. To encode m, the 'level 1' encoding algorithm computes $c = \mathsf{Enc}_1(m) = \sum_i h_i s_i + m \in R_q$, with s_i small from χ_ρ. Note that c has the form $c = pg'/f + m_0$, for small g' which is essentially an NTRU ciphertext of $m_0 = m$ mod p, with no extra error terms pe_i, as in the original NTRUEncrypt scheme. The add and mul algorithms just perform addition and multiplication over R_q as in the basic homomorphic encryption scheme of the previous section. As a consequence, level k encodings of m_0 have the form $c = pg'/f^k + m_0$ for small g'. For $x \in R_q$, let $MSB_\ell(x)$ denote the polynomial whose coefficients consist of the most-significant ℓ bits of each coefficient of x. Given such a level k encoding c, the extraction algorithm Ext computes the representative $r(m_0) = \mathsf{Ext}(\{h_i\}_i, e_k) = MSB_\ell(e_k \cdot c) = MSB_\ell(ug' + uf^k/pm_0) = MSB_\ell(ug'' + u/pm_0)$ for some small g''. Note that ug'' is an element of 'small' norm $O(\text{poly}(n)^k \cdot q^{1/2})$ compared with q, if q is chosen sufficiently larger than $\text{poly}(n)^{2k}$, whereas $u/pm_0 \in R_q$ is a large element due to the large p^{-1} factor and the fact that p does not divide hm_0 in R with overwhelming probability. This means that the ℓ most-significant bits (when ℓ is chosen smaller than $\log q - \log(\|ug''\|)$) of $r(m_0)$ are, with high probability, determined only by the large u/pm_0 term, that is dependent only on the encoded element m_0 (and the fixed elements u, p) and not on any randomness in the encoding, as required.

We conclude this section by stating the NTRU k-graded discrete-log problem that is necessary for the security of the Diffie-Hellman key exchange based on the above problem (actually the security of the k-graded Decision Diffie-Hellman is necessary, but even the hardness of the simpler k-graded discrete-log problem is open).

k-graded NTRU Discrete-Log Problem $\texttt{DNDL}_{k,n,q,\phi,\chi_\eta,\chi_\sigma,\chi_\beta,\chi_\rho,\ell}$**.** Given (k, n, q, ϕ), $h_1 = pg_1/f, \ldots, h_{m_r} = pg_{m_r}/f$, $e_k = uf^k/p \in R_q$ and $c = \sum_i h_i s_i + m \in R_q$, with p sampled from $\chi_\eta = D_{\mathbb{Z}^n,\eta}$ with $\eta = \text{poly}(n)$, $g_1 \ldots, g_{m_r}, f, m$ sampled from the 'small' distribution χ_σ subject to $f = 1$ mod p with $\sigma = \text{poly}(n) \cdot \|p\|$, $s_1, \ldots, s_{m_r}$ sampled from the 'small' distribution χ_β with $\beta = \text{poly}(n) \cdot p$, and u sampled from the 'small' distribution χ_ρ with $\rho = q^{1/2}\text{poly}(n)$, find m' with $\|m'\|$ 'small' (less than q) such that $m' = m$ mod p.

Note that ignoring e_k, this problem is a variant of the NTRU ciphertext cracking problem with a secret p. But the presence of e_k makes this problem quite different;

[1] Some desirable properties can be established for $m_r = 2$ (see [40] for more details).

indeed knowledge of e_k allows an element m' in coset of m modulo p to be efficiently recovered. In particular, the attacker can compute several quantities of the form $v_r = e_k \cdot (\prod_{j=1,\ldots,s} \sum_{i \leq m_r} h_i \rho_{i,j}) \cdot (1+h_1)^{k-s}$ for 'small' random $\rho_{i,j}$. Since $\sum_{i \leq m_r} h_i \rho_{i,j}$ has the form $(p \cdot g'_j)/f$ with 'small' random g'_j with $j = 1, \ldots, s$ while $1 + h_1 = (f + pg_1)/f = (1 + pg''_1)/f$ with small g''_1, it follows that that the v_r are of the form $v_r = u\hat{g}_r$ if $s = 1$ and $v_r = up\hat{g}_r$ if $s \geq 2$ for some small random $\hat{g}_r$, where the equality holds over R, not just over R_q. From the latter multiples of u and up, one can typically compute efficiently a (large norm) basis for the ideal (p) generated by p (see Sec.6.3.3 in [23]). Then, given the encoding $c = \sum_i h_i s_i + m \in R_q$, one could compute $c' = e_k \cdot h_1 \cdot c \cdot (1 + h_1)^{k-2} = ug_1 \cdot (m + tp) \in R$ for some 'small' t and similarly $c'' = e_k \cdot h_1 \cdot (1+h_1)^{k-1} = ug_1 \cdot (1+t'p) \in R$ for some 'small' t'. Note that $c' = ug_1 m \bmod p$ and $c'' = ug_1 \bmod p$, so that $m' = c' \cdot (c'')^{-1} \bmod p = m \bmod p$, assuming that ug_1 invertible in R_p. However, since the attacker's basis for (p) has large norm, the computed element m' also has large norm, and, due to the apparent difficulty of computing a *short* basis for the ideal (p), it still seems hard to efficiently compute a *short* (norm less than q) representative (such as m) of the coset of m' modulo (p). We remark that other applications require the hardness of yet other interesting variants of this problem. We refer the interested reader to [23, 40] as well as [24, 8], where a variant of the GGH construction is applied to provide a candidate solution to the problem of *cryptographic program obfuscation* which seems very promising in terms of its potential applications (see [66, 11, 3] for some examples).

6 Conclusions

We surveyed recent developments in both the security analysis and applications of the NTRU cryptosystem and its variants. Some of these developments motivate the study of new computational problems on polynomial rings, whereas others help to unify the field of lattice-based cryptography, by showing that the security of the NTRU system can be based on the same foundations as more recent lattice-based schemes. The simplicity of the NTRU system seems to give it potential efficiency advantages compared with other known lattice-based systems (e.g., unlike other known public-key encryption schemes based on the Ring-LWE problem, NTRU ciphertexts consist of just a *single* ring element), and also seems to make it easier to construct some powerful cryptographic functionalities, such as the multilinear maps discussed in Section 5.2 and the *multikey* homomorphic properties of the FHE scheme in Section 5.1. Yet the full fundamental potential of the NTRU system, as well as that of other lattice-based systems, is not completely clear at present. We hope this survey will encourage more progress in this active field in the years to come.

Acknowledgments. The author would like to thank Damien Stehlé and Igor Shparlinksi for useful discussions on the material presented here, and the anonymous referees for their helpful comments.

Bibliography

[1] M. Ajtai, Generating hard instances of lattice problems (extended abstract), in: *Proc. of STOC*, pp. 99–108, ACM, 1996.

[2] M. Ajtai and C. Dwork, A public-key cryptosystem with worst-case/average-case equivalence, in: *Proceedings of the 29th Symposium on the Theory of Computing (STOC 1997)*, pp. 284–293, ACM, 1997.

[3] P. Ananth, D. Boneh, S. Garg, A. Sahai and M. Zhandry, Differing-Inputs Obfuscation and Applications, *IACR Cryptology ePrint Archive* 2013 (2013), 689, `http://eprint.iacr.org/2013/689`.

[4] B. Applebaum, D. Cash, C. Peikert and A. Sahai, Fast Cryptographic Primitives and Circular-Secure Encryption Based on Hard Learning Problems, in: *Proc. of CRYPTO*, LNCS 5677, pp. 595–618, Springer, 2009.

[5] L. Babai, On Lovász lattice reduction and the nearest lattice point problem, *Combinatorica* 6 (1986), 1–13.

[6] W. Banaszczyk, New bounds in some transference theorems in the geometry of numbers, *Math. Ann* 296 (1993), 625–635.

[7] W.D. Banks and I.E. Shparlinski, Distribution of Inverses in Polynomial Rings, *Indag. Math.* 12 (2001), 303–315.

[8] B. Barak, S. Garg, Y. Tauman Kalai, O. Paneth and A. Sahai, Protecting Obfuscation Against Algebraic Attacks, *IACR Cryptology ePrint Archive* 2013 (2013), 631, To appear at Eurocrypt 2014. `http://eprint.iacr.org/2013/631`.

[9] D. Boneh and M. Franklin, Identity-based encryption from the Weil pairing, *SIAM J. Comput* 32 (2003), 586–615.

[10] D. Boneh and A. Silverberg, Applications of Multilinear Forms to Cryptography, *IACR Cryptology ePrint Archive, Report 2002/080* 2002 (2002).

[11] D. Boneh and M. Zhandry, Multiparty Key Exchange, Efficient Traitor Tracing, and More from Indistinguishability Obfuscation, *IACR Cryptology ePrint Archive* 2013 (2013), 642, `http://eprint.iacr.org/2013/642`.

[12] J.W. Bos, K. Lauter, J. Loftus and M. Naehirg, Improved Security for a Ring-Based Fully Homomorphic Encryption Scheme, in: *Cryptography and Coding*, pp. 45–64, 2013.

[13] Z. Brakerski, Fully Homomorphic Encryption without Modulus Switching from Classical GapSVP, in: *CRYPTO*, pp. 868–886, 2012.

[14] Z. Brakerski, C. Gentry and V. Vaikuntanathan, (Leveled) fully homomorphic encryption without bootstrapping, in: *ITCS*, pp. 309–325, 2012.

[15] Z. Brakerski, A. Langlois, C. Peikert, O. Regev and D. Stehlé, Classical hardness of learning with errors, in: *STOC*, pp. 575–584, 2013.

[16] Z. Brakerski and V. Vaikuntanathan, Efficient Fully Homomorphic Encryption from (Standard) LWE, in: *FOCS*, pp. 97–106, 2011.

[17] _____, *Lattice-Based FHE as Secure as PKE*, IACR Cryptology ePrint Archive, Report 2013/541, 2013.

[18] J. Buchmann, D. Cabarcas, F. Göpfert, A. Hülsing and P. Weiden, Discrete Ziggurat: A Time-Memory Trade-off for Sampling from a Gaussian Distribution over the Integers, *IACR Cryptology ePrint Archive* 2013 (2013), 510, To appear at SAC 2013. `http://eprint.iacr.org/2013/510`.

[19] D. Coppersmith and A. Shamir, Lattice Attacks on NTRU, in: *Proc. of Eurocrypt*, LNCS 1233, pp. 52–61, Springer, 1997.

[20] W. Diffie and M. E. Hellman, New Directions in Cryptography, *IEEE Trans. Inform. Theory* IT-22 (1976), 644–654.

[21] L. Ducas, A. Durmus, T. Lepoint and V. Lyubashevsky, Lattice Signatures and Bimodal Gaussians, in: *CRYPTO (1)*, pp. 40–56, 2013.

[22] L. Ducas and P. Q. Nguyen, Faster Gaussian Lattice Sampling Using Lazy Floating-Point Arithmetic, in: *ASIACRYPT*, pp. 415–432, 2012.

[23] S. Garg, C. Gentry and S. Halevi, Candidate Multilinear Maps from Ideal Lattices, in: *EUROCRYPT*, pp. 1–17, 2013.

[24] S. Garg, C. Gentry, S. Halevi, M. Raykova, A. Sahai and B. Waters, Candidate Indistinguishability Obfuscation and Functional Encryption for all Circuits, in: *FOCS*, pp. 40–49, 2013.

[25] C. Gentry, Fully homomorphic encryption using ideal lattices, in: *Proc. of STOC*, pp. 169–178, ACM, 2009.

[26] C. Gentry, J. Jonsson, J. Stern and M. Szydlo, Cryptanalysis of the NTRU Signature Scheme (NSS) from Eurocrypt 2001, in: *Proc. of Asiacrypt*, LNCS 2248, pp. 1–20, Springer, 2001.

[27] C. Gentry, C. Peikert and V. Vaikuntanathan, Trapdoors for hard lattices and new cryptographic constructions, in: *Proc. of STOC*, pp. 197–206, ACM, 2008, Full version available at `http://eprint.iacr.org/2007/432.pdf`.

[28] C. Gentry, A. Sahai and B. Waters, Homomorphic Encryption from Learning with Errors: Conceptually-Simpler, Asymptotically-Faster, Attribute-Based, in: *CRYPTO (1)*, pp. 75–92, 2013.

[29] C. Gentry and M. Szydlo, Cryptanalysis of the Revised NTRU Signature Scheme, in: *Proc. of Eurocrypt*, LNCS 2332, pp. 299–320, Springer, 2002.

[30] S. Goldwasser and S. Micali, Probabilistic Encryption, *Journal of Computer and System Sciences* 28 (1984), 270–299.

[31] J. Hoffstein, N. Howgrave-Graham, J. Pipher and W. Whyte, *Practical lattice-based cryptography:* `NTRUEncrypt` *and* `NTRUSign`, 2009, Chapter of [55].

[32] J. Hoffstein, J. Pipher and J. H. Silverman, *NTRU: a new high speed public key cryptosystem*, Preprint; presented at the rump session of Crypto'96, 1996.

[33] _____, NTRU: a ring based public key cryptosystem, in: *Proc. of ANTS*, LNCS 1423, pp. 267–288, Springer, 1998.

[34] J. Hoffstein and J. H. Silverman, Optimizations for NTRU, in: *Public-Key Cryptography and Computational Number Theory (Warsaw, Sep. 11-15, 2000)*, 2000, Available at `http://www.securityinnovation.com/uploads/Crypto/TECH_ARTICLE_OPT.pdf`.

[35] N. Howgrave-Graham, P. Q. Nguyen, D. Pointcheval, J. Proos, J. H. Silverman, A. Singer and W. Whyte, The Impact of Decryption Failures on the Security of NTRU Encryption, in: *Proc. of CRYPTO*, LNCS 2729, pp. 226–246, Springer, 2003.

[36] IEEE P1363, *Standard Specifications For Public-Key Cryptography*, `http://grouper.ieee.org/groups/1363/`.

[37] K. Jarvis and M. Nevins, ETRU: NTRU over the Eisenstein integers, *Designs, Codes and Cryptography* (2013).

[38] A. Joux, A One Round Protocol for Tripartite Diffie-Hellman, in: *ANTS*, pp. 385–394, 2000.

[39] J. Katz and Y. Lindell, *Introduction to Modern Cryptography*, Cryptography and Network Security, Chapman and Hall/CRC Press, 2008.

[40] A. Langlois, D. Stehlé and R. Steinfeld, GGHLite: More Efficient Multilinear Maps from Ideal Lattices, in: *EUROCRYPT*, 2014, To appear.

[41] A. K. Lenstra, H. W. Lenstra, Jr. and L. Lovász, Factoring polynomials with rational coefficients, *Math. Ann* 261 (1982), 515–534.

[42] L. Li and O. Roche-Newton, An Improved Sum-Product Estimate for General Finite Fields, *arXiv.org Archive* arXiv:1106.1148 [math.CO] (2011).

[43] A. López-Alt, E. Tromer and V. Vaikuntanathan, On-the-fly multiparty computation on the cloud via multikey fully homomorphic encryption, in: *Proc. of STOC*, pp. 1219–1234, 2012.

[44] V. Lyubashevsky, Fiat-Shamir with Aborts: Applications to Lattice and Factoring-Based Signatures, in: *Proc. of ASIACRYPT*, LNCS 5912, pp. 598–616, Springer, 2009.

[45] V. Lyubashevsky, Lattice Signatures without Trapdoors, in: *EUROCRYPT*, pp. 738–755, 2012.

[46] V. Lyubashevsky and D. Micciancio, Generalized Compact Knapsacks Are Collision Resistant, in: *Proc. of ICALP*, LNCS 4052, pp. 144–155, Springer, 2006.

[47] V. Lyubashevsky, C. Peikert and O. Regev, On Ideal Lattices and Learning with Errors over Rings, in: *Proc. of EUROCRYPT*, LNCS 6110, pp. 1–23, Springer, 2010.

[48] ———, On Ideal Lattices and Learning with Errors over Rings, *J. ACM* 60 (2013).

[49] D. Micciancio, Generalized Compact Knapsacks, Cyclic Lattices, and Efficient One-Way Functions, *Comput. Complexity* 16 (2007), 365–411.

[50] D. Micciancio and O. Regev, Worst-case to Average-case Reductions based on Gaussian Measures, *SIAM J. Comput* 37 (2007), 267–302.

[51] ———, Lattice-based cryptography, in: *Post-Quantum Cryptography, D. J. Bernstein, J. Buchmann, E. Dahmen (Eds)*, pp. 147–191, Springer, 2009.

[52] D. Micciancio and P. Voulgaris, A deterministic single exponential time algorithm for most lattice problems based on Voronoi cell computations, in: *Proc. of STOC*, pp. 351–358, ACM, 2010.

[53] S. Min, G. Yamamoto and K. Kim, Weak Property of Malleability in NTRUSign, in: *Proc. of ACISP*, LNCS 3108, pp. 379–390, Springer, 2004.

[54] P. Q. Nguyen and O. Regev, Learning a Parallelepiped: Cryptanalysis of GGH and NTRU Signatures, *Journal of Cryptology* 22 (2009), 139–160.

[55] P. Q. Nguyen and B. Vallée (editors), *The LLL Algorithm: Survey and Applications*, Information Security and Cryptography, Springer, 2009.

[56] C. Peikert, Limits on the Hardness of Lattice Problems in ℓ_p Norms, *Comput. Complexity* 2 (2008), 300–351.

[57] C. Peikert and A. Rosen, Efficient Collision-Resistant Hashing from Worst-Case Assumptions on Cyclic Lattices, in: *Proc. of TCC*, LNCS 3876, pp. 145–166, Springer, 2006.

[58] R. A. Perlner and D. A. Cooper, Quantum resistant public key cryptography: a survey, in: *Proc. of IDtrust*, pp. 85–93, ACM, 2009.

[59] O. Regev, On lattices, learning with errors, random linear codes, and cryptography, in: *Proc. of STOC*, pp. 84–93, ACM, 2005.

[60] ———, On lattices, learning with errors, random linear codes, and cryptography, in: *Proc. of STOC*, pp. 84–93, ACM, 2005.

[61] ———, On lattices, learning with errors, random linear codes, and cryptography, *J. ACM* 56 (2009).

[62] ———, *The Learning with Errors Problem*, 2010, Invited survey in CCC 2010, available at `http://www.cs.tau.ac.il/~odedr/`.

[63] R. L. Rivest, A. Shamir and L. M. Adleman, A Method for Obtaining Digital Signatures and Public-Key Cryptosystems, *Communications of the ACM* 21 (1978), 120–126.

[64] R.L. Rivest, L. Adelman and M.L. Detouzos, On Databanks and Privacy Homomorphisms, in: *Foundations of Secure Computations*, Academic Press, 1978.

[65] S. Sinha Roy, F. Vercauteren and I. Verbauwhede, High Precision Discrete Gaussian Sampling on FPGAs, To appear at SAC 2013. Available at http://www.cosic.esat.kuleuven.be/publications/article-2372.pdf.

[66] A. Sahai and B. Waters, How to Use Indistinguishability Obfuscation: Deniable Encryption, and More, *IACR Cryptology ePrint Archive* 2013 (2013), 454, `http://eprint.iacr.org/2013/454`.

[67] R. Sakai, K. Ohgishi and M. Kasahara, Cryptosystems based on pairing, in: *SCIS*, 2000.

[68] C. P. Schnorr, A Hierarchy of Polynomial Lattice Basis Reduction Algorithms, *Theor. Comput. Science* 53 (1987), 201–224.

[69] C.-P. Schnorr, Efficient Signature Generation by Smart Cards, *Journal of Cryptology* 4 (1991), 161–174.

[70] P. W. Shor, Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer, *SIAM Review* 41 (1999), 303–332.

[71] D. Stehlé and R. Steinfeld, Making NTRU as Secure as Worst-Case Problems over Ideal Lattices, in: *Proc. of EUROCRYPT*, pp. 27–47, 2011.

[72] _____, Making NTRUEncrypt and NTRUSign as Secure as Standard Worst-Case Problems over Ideal Lattices, *IACR Cryptology ePrint Archive* 2013 (2013), 004, Extended version of Eurocrypt 2011 paper. `http://eprint.iacr.org/2013/004`.

[73] D. Stehlé, R. Steinfeld, K. Tanaka and K. Xagawa, Efficient Public Key Encryption Based on Ideal Lattices, in: *Proc. of ASIACRYPT*, LNCS 5912, pp. 617–635, Springer, 2009.

[74] M. Szydlo, Hypercubic Lattice Reduction and Analysis of GGH and NTRU Signatures, in: *Proc. of Eurocrypt*, LNCS 2656, pp. 433–448, Springer, 2003.

Author information

Ron Steinfeld, Clayton School of IT, Faculty of IT, Monash University, Australia, Australia.
E-mail: `ron.steinfeld@monash.edu`