Concrete Analysis of Approximate Ideal-SIVP to Decision Ring-LWE Reduction

Neal Koblitz
Department of Mathematics, Box 354350,
University of Washington, Seattle, WA 98195 U.S.A.
email: koblitz@uw.edu

Subhabrata Samajder
Department of Computer Science and Engineering,
Indraprastha Institute of Information Technology, Delhi, 110020 India
email: subhabrata@iiitd.ac.in

Palash Sarkar and Subhadip Singha
Applied Statistics Unit,
Indian Statistical Institute, Kolkata, India 700108
email: {palash, subha_r}@isical.ac.in

April 19, 2022

Abstract

A seminal 2013 paper by Lyubashevsky, Peikert, and Regev proposed basing post-quantum cryptography on ideal lattices and supported this proposal by giving a polynomial-time security reduction from the approximate Shortest Independent Vectors Problem (SIVP) to the Decision Learning With Errors (DLWE) problem in ideal lattices. We give a concrete analysis of this multi-step reduction. We find that the tightness gap in the reduction is so great as to vitiate any meaningful security guarantee, and we find reasons to doubt the feasibility in the foreseeable future of the quantum part of the reduction. In addition, when we make the reduction concrete it appears that the approximation factor in the SIVP problem is far larger than expected, a circumstance that causes the corresponding approximate-SIVP problem most likely not to be hard for proposed cryptosystem parameters. We also discuss implications for systems such as Kyber and SABER that are based on module-DLWE.

Keywords: ideal lattices, shortest vector problem, ring learning with errors, concrete analysis.

Mathematics Subject Classification: 94A60

Contents

1 Introduction

In 2013 Lyubashevsky, Peikert, and Regev [21] published a security reduction in support of proposed post-quantum cryptography based on the difficulty of the Decision Learning With Errors (DLWE) problem in a lattice coming from the embeddings of an ideal of a number field[1]. Their elaborate, multi-step reduction showed that the worst-case γ-approximate Shortest Independent Vectors Problem (SIVP$_\gamma$) for ideal lattices could be solved with polynomially many calls to an oracle that solves DLWE for ideal lattices. Our purpose in this paper is to analyze this reduction in concrete terms.

The U.S. government's NIST is currently running a multi-year competition[2] to select candidates for standardisation of post-quantum public key cryptography. Some of the proposals under consideration are based on lattices. Two of the lattice-based finalists, namely Kyber [7] and SABER [12], are based on module lattices [20, 27] which are generalisations of ideal lattices. In [27] Peikert and Pepin show that ring-LWE reduces to module-LWE for a given size of the module and ring and then argue for the security of module-LWE-based systems by citing presumed hardness of ring-LWE. In addition, in [20] Langlois and Stehlé give a reduction from approximate module-SIVP to module-DLWE. In §7.2 we discuss the generalisation from ring-DLWE to module-DLWE.

1.1 The structure of the approximate ideal-SIVP to decision ring-LWE reduction

The structure of the reduction in [21] – a nested sequence of intermediate reductions – gives rise to two difficulties from a practice-oriented perspective. In the first place, the tightness gaps multiply from one reduction to the next. If algorithm A calls on algorithm B m times, and B calls on C n times, then there are mn calls on C. We found that the cumulative tightness gap in the reduction is so great as to render the security guarantee meaningless for practical parameter values.

In the second place, seven of the nested reductions take place within a quantum computer. We have reasons to doubt the feasibility of the quantum part of the reduction even assuming the advent of quantum computers that are scaled to much larger size than what is needed to break RSA and ECC.

1.2 Restricting to a special class of lattices

From the beginning, the strongest argument advanced for lattice-based cryptography has been worst-to-average case reductions. The hardest instances of problems such as the approximate Shortest Vector Problem (SVP) and the approximate Shortest Independent Vectors Problem (SIVP) can sometimes be shown to reduce to a

[1] Earlier D. Stehlé *et al.* [37] had published a security reduction for ideal lattices, but their reduction only goes as far as search ring-LWE, not ring-DLWE. The part of the reduction from LWE to DLWE was carried out in [21] only for cyclotomic number fields generated by power-of-2 roots of unity; in [13] this reduction was extended to all cyclotomic number fields.

[2] https://csrc.nist.gov/Projects/post-quantum-cryptography/round-3-submissions, accessed on February 8, 2022.

random instance of the lattice problem that the proposed cryptosystem is based on. This argument loses validity if lattices are chosen from a special class, such as the class of ideal lattices, rather than from the set of all general lattices, unless one has evidence that the shortest vector problems and their approximate variants do not lose any of their worst-case intractability when restricted to the special class of ideal lattices.

As far as we are aware, no such evidence exists. After arguing for many years that worst-to-average case reductions are important in order to have confidence in security, it seems that some promoters of lattice-based systems have undermined that argument by changing course and now preferring to work in a special subclass of lattices.

From a number theory perspective, one reason to wonder about the effect of specializing to ideal lattices is that they have much more structure than general lattices, notably the presence of isomorphisms between different embeddings of the number field. In the cyclotomic case (and, more generally, in the case of non-cyclotomic Galois fields as well) the isomorphisms are all automorphisms (which permute the roots of unity), and those automorphisms were in fact used to good effect in the security reductions in [21].

But the much greater structure and symmetry, especially in the cyclotomic case, also make it likely that the supposedly hard shortest vector problems that are the basis for the security of lattice-based protocols are in fact much easier in this restricted setting. For example, as pointed out in [21], in cyclotomic ideal lattices SIVP_γ is trivially equivalent to SVP_γ, because a short vector can be multiplied by roots of unity to get a full set of independent vectors of the same length. In contrast, in the general case only a much weaker result is known, namely, that $\mathrm{SIVP}_{\sqrt{n}\gamma}$ reduces in polynomial time to SVP_γ, where n is the dimension of the lattice [22]. This suggests that for general lattices SIVP and SIVP_γ are strictly harder than SVP and SVP_γ – a separation that disappears when restricted to cyclotomic ideal lattices. Thus, even if SVP and approximate SVP for cyclotomic ideal lattices were to be as hard as for general lattices, SIVP and approximate SIVP for cyclotomic ideal lattices would likely be easier.

In the simplest case of cyclotomic fields generated by m-th roots of unity with $m = 3, 4$, i.e., with $n = \varphi(m) = 2$, SVP/SIVP is trivial for any ideal lattice[3], whereas the general SVP/SIVP in two dimensions is not completely trivial. It is not yet clear whether this gap in difficulty increases in higher dimensions, but there is no reason to assume that it does not. Moreover, for the full lattice R, where R is the ring of integers of a cyclotomic field, Lemma 2.9 of [21] tells us that the ℓ_2-norm of a shortest vector is $\geq \sqrt{n}$. Since $\|1\| = \sqrt{n}$, that means that 1 is a shortest vector. If the ideal $\mathcal{I}$ is principal (which is true for all ideals for $n = 2, 4, 8, 16$), then for each imbedding σ the image of $\mathcal{I}$ is just a scaled (and rotated) version of the image of R. That does not immediately lead to a simple result for the shortest vector in $\mathcal{I}$, but it certainly suggests a close relationship between the geometry of $\mathcal{I}$ and the geometry of the unit ideal lattice R, for which SVP and SIVP are trivial.

1.3 The role of the approximation factor

Another reason why the reduction of SIVP_γ to SVP_γ for cyclotomic ideal lattices is troubling is that for $\gamma > n$, Goldreich and Goldwasser [15] showed that SVP_γ is unlikely to be NP-hard. In [21] the problem assumed to be hard is SIVP_γ where $\gamma = \tilde{O}(\sqrt{n}/\alpha)$ in which $\alpha < (\ln n/n)^{1/2}$ is much less than 1. The Goldreich-Goldwasser result shows that this approximate SIVP problem for cyclotomic ideal lattices is almost certainly not NP-hard. Further, the GapSVP problem on ideal lattices is easy [32], while it is conjectured to be hard for general lattices.

The last sentence in the statement of the "main theorem" (Theorem 3.6) of [21] says that the target problem for the reduction from $\tilde{O}(\sqrt{n}/\alpha)$-approximate SIVP can be taken to be decision ring-$\mathrm{LWE}_{q,D_{r_0}}$ with

$$r_0 = \alpha(n\ell/\log(n\ell))^{1/4}, \tag{1}$$

where ℓ is the number of queries made by the DLWE-distinguisher. Based on the estimate for α in the previous

[3]The case $m = 3$ corresponds to tiling of the plane using equilateral triangles, whereas the case $m = 4$ corresponds to tiling the plane with squares.

paragraph, this suggests a $\tilde{O}(n)$ approximation factor for the short independent vectors problems that are the basis for the security of DLWE-based protocols. This appears to be incorrect.

First of all, cryptosystems based on the decision-LWE problem generally have a fixed and publicly known Gaussian distribution width r_0, which is assumed to be roughly of order $n^{-1/2}$. In that case, ignoring the contribution of ℓ, from (1) we find a larger approximation factor $\tilde{O}(n^{5/4})$.

In the second place, when filling in the details of the reduction, we found that in place of (1) we needed

$$r_0 = \alpha(N_2 n\ell/\log(N_2 n\ell))^{1/4}, \tag{2}$$

where N_2 is a parameter for the reduction algorithm that is of order at least $n^2\delta_2^{-2}$. This causes the approximation factor in the supposedly intractable SIVP to be greater than $n^{7/4}\delta_2^{-1/2}$. There is little reason to have confidence in the intractability of $\mathrm{SIVP}_{n^{7/4}\delta_2^{-1/2}}$ for cyclotomic ideal lattices.

1.4 Efficiency versus security

Increased efficiency of implementation is the main reason for specializing lattice-based cryptography to the lattices coming from ideals of number rings and, in particular, to cyclotomic ideal lattices. As explained in [21], general lattices "are rather inefficient due to an inherent quadratic overhead in the use of LWE," whereas ideal lattices provide a major speed-up in running time and a reduction in the size of public keys by a factor of n, the dimension of the lattice.

The abstract of [21] also uses the same term *efficient* in a very different sense of the word in describing the security reductions in their paper. When they speak of an "efficient security reduction" they mean a polynomial-time quantum reduction. The question of practical feasibility is not addressed in the paper.

This is an unfortunate omission. Ever since Bellare, Rogaway, and others argued for "practice-oriented provable security" in the 1990s [3], it has been widely recognized that a close examination of security reductions to determine the real-world guarantees that they give is essential. Moreover, it is a cardinal principle of cryptography that efficiency of usage should not be prioritized over meaningful evidence of security.

In nearly a decade since the appearance of [21] several attacks on ideal lattice problems have confirmed the intuition that such lattices are more vulnerable to attacks, both classical and quantum, than general lattices [11, 5]. These works suggest that from a security standpoint cryptography based on ring-LWE and similar ideal lattice problems might not stand the test of time.

1.5 Related reductions

In [31] Regev introduced the LWE problem and gave a reduction from approximate SIVP over general lattices to Decision LWE. This work is generally considered a breakthrough in lattice-based cryptography and spurred a great deal of subsequent research. The concrete aspect of the reduction in [31] was analysed in [10] and the analysis was refined in [35]. Commenting on the concrete analysis of [31] in [10], Bernstein [4] remarked that "the loss of tightness is gigantic." A different concrete analysis of the reduction in [31] was carried out in [14] and this work also considered increasing the value of n to compensate for the tightness gap. We revisit the tightness gap of the reduction in [31] and obtain a more accurate estimate. We point out several aspects that were overlooked or were overestimated in [10, 35, 14]. The resulting gap turns out to be greater than the previous estimates. Our present work shows that the tightness gap in the reduction in [21] is even greater. Further, we argue that trying to increase the value of n to compensate for the tightness gap in [31, 21] is not a meaningful exercise because of the very large values of n that would be needed.

The reduction in [31] is quantum. Later work pursued the goal of obtaining a classical reduction. The first result in this direction was obtained by Peikert [25], who gave a classical reduction from GapSVP over general lattices to the LWE problem. The drawback of this result was that it required an exponential size modulus. A subsequent work by Brakerski *et al.* [8] gave a reduction where the modulus is of polynomial size. This

reduction was also unsatisfactory since it reduced GapSVP on a lattice of dimension $\sqrt{n}$ to the n-dimensional LWE problem. Since GapSVP is not hard over ideal lattices, the approach adopted in [25, 8] is not meaningful for such lattices. A concrete analysis of the reductions in [25, 8] was carried out in [34].

1.6 The danger of relying on an earlier reduction argument

A potential problem arises when a series of authors rely upon an earlier security reduction to form part of new proofs of security, essentially regarding the former as a black box. They then inherit any tightness gap, poorly grounded hardness assumption, or unrealistic assumption about feasibility that might be in the earlier reduction.

We know of at least six papers that rely upon the quantum reduction in [31] to obtain their results: the reduction from ideal-SIVP to the problem of breaking the Stehlé-Steinfeld version of NTRU [36], the reductions from ideal-SIVP to ring-DLWE for cyclotomic fields [21] and for general number fields [28], and the reduction from module-SIVP to module-DLWE [20]. Other papers that use the proofs in [21] (and hence also in [31]) as a black box include [13, 39]. Any doubts that arise about the quantum part of the original reduction in [31] then also apply to all of the later papers as well.

Ideally, authors should carefully examine both the explicit and implicit assumptions being made in an earlier work before incorporating the earlier reduction into their own security proofs. If this is not done, one risks having cascading assumptions that eventually resemble a house of cards.

1.7 Outline of the paper

The background and preliminaries required for the paper are given in §2. The reduction in [21] can be divided into two parts. The first part is a reduction from approximate ideal-SIVP to the search ring-LWE problem, while the second part is a reduction from the search ring-LWE problem to the decision ring-LWE problem. The concrete analysis of the first part is described in §3 and that for the second part is described in §4. The two parts are combined and the end-to-end reduction from approximate ideal-SIVP to decision ring-LWE is summarised in §5. A detailed discussion of several problematic issues with the quantum aspect of the reduction is given in §6. The reduction in [21] holds for cyclotomic number fields. A follow-up work [29] gave a reduction from approximate ideal-SIVP to decision ring-LWE for any number field. In §7 we perform a concrete analysis of the reduction in [29] and show that the tightness gap is much larger than the tightness gaps of either [31] or [21]. Section 8 has concluding remarks. Some important mathematical details of the reduction are presented in Appendices A and B.

2 Preliminaries

A brief summary of the relevant concepts is provided below. For further details the reader may refer to [21]. We note though that at certain places we have simplified the description that is in [21].

By $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ we will denote the sets of integers, rationals, reals and complex numbers respectively. Let n be a positive integer. For a vector $\mathbf{a} = (a_1, \ldots, a_n)$ in $\mathbb{R}^n$ or $\mathbb{C}^n$, the ℓ_2-norm of $\mathbf{a}$ is defined to be $\|\mathbf{a}\| = (|a_1|^2 + \cdots + |a_n|^2)^{1/2}$ and the ℓ_∞-norm of $\mathbf{a}$ is defined to be $\|\mathbf{a}\|_\infty = \max_{i\in[n]} |a_i|$. We will mostly consider the ℓ_2-norm. At a few places, the ℓ_∞-norm is used and we will explicitly identify these cases.

Let s_1 and s_2 be non-negative integers such that $s_1 + 2s_2 = n$. The space $H \subset \mathbb{R}^{s_1} \times \mathbb{C}^{2s_2}$ is defined as

$$H \quad = \quad \{(x_1, \ldots, x_n) \in \mathbb{R}^{s_1} \times \mathbb{C}^{2s_2} : x_{s_1+s_2+j} = \overline{x_{s_1+j}}, j = 1, \ldots, s_2\}. \tag{3}$$

Using the inner product on H induced on it by $\mathbb{C}^n$, it can be shown that H is isomorphic to $\mathbb{R}^n$ as an inner product space. For $j \in [n]$, let $\mathbf{e}_j \in \mathbb{C}^n$ be the vector which has 1 in its j-th component and 0 elsewhere. An orthonormal basis for H is given by $\{\mathbf{h}_i\}_{i\in[n]}$, where for $j \in [s_1]$, $\mathbf{h}_j = \mathbf{e}_j$ and for $s_1 < j \leq s_1 + s_2$, $\mathbf{h}_j = (\mathbf{e}_j + \mathbf{e}_{j+s_2})/\sqrt{2}$,

$\mathbf{h}_{j+s_2} = \sqrt{-1}(\mathbf{e}_j - \mathbf{e}_{j+s_2})/\sqrt{2}$. When $\mathbf{x} \in H$ is written in terms of the orthonormal basis as $\mathbf{x} = \sum_{i=1}^{n} a_i \mathbf{h}_i$ with $(a_1, ..., a_n) \in \mathbb{R}^n$, the norm of $\mathbf{x}$ is simply $\|(a_1, \ldots, a_n)\|$.

A lattice is a discrete additive subgroup of H. Let $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\} \subset H$ be a set of linearly independent vectors. The full rank lattice generated by B is defined to be $\mathcal{L}(B) = \{\sum_{i=1}^{n} z_i \mathbf{b}_i : (z_1, \ldots, z_n) \in \mathbb{Z}^n\}$. Given a lattice basis B, the fundamental parallelepiped $\mathcal{P}(B)$ is defined to be the set $\{B\mathbf{x} : \mathbf{x} \in \mathbb{R}^n,\ 0 \leq x_i < 1\}$. For an n-dimensional lattice Λ, let $\lambda_i(\Lambda)$ with $i \in \{1, \ldots, n\}$, be the least real number r such that Λ has i linearly independent vectors with the longest having length r with respect to the ℓ_2-norm. In particular, $\lambda_1(\Lambda) = \min_{\mathbf{x} \in \Lambda \setminus \{0\}} \|\mathbf{x}\|$ is called the minimum distance of the lattice. The dual of a lattice Λ is defined to be $\Lambda^* = \{\mathbf{x} \in H : \langle \mathbf{y}, \mathbf{x} \rangle \in \mathbb{Z}, \text{ for all } \mathbf{y} \in \Lambda\}$, where $\langle \mathbf{y}, \mathbf{x} \rangle = \sum y_i \overline{x}_i$ is the inner product. Theorem 2.1 of [2] shows that for any n-dimensional lattice Λ, $1 \leq \lambda_1(\Lambda) \cdot \lambda_n(\Lambda^*) \leq n$.

For $r > 0$, the Gaussian function $\rho_r : H \to (0, 1]$ is defined to be $\rho_r(\mathbf{x}) = \exp(-\pi \|\mathbf{x}\|^2 / r^2)$. The continuous Gaussian probability distribution D_r over H is given by the density function $r^{-n}\rho_r(\mathbf{x})$. Note that D_r is the n-dimensional normal distribution with mean vector $(0, \ldots, 0)$ and variance/co-variance matrix $\mathrm{diag}(\sigma^2, \ldots, \sigma^2)$ where $\sigma = r/\sqrt{2\pi}$. Consequently, if X_1 and X_2 are independent random variables following D_{r_1} and D_{r_2}, then $X_1 + X_2$ follows $D_{\sqrt{r_1^2 + r_2^2}}$ (see Theorem 4.2.14 of [9]).

For a lattice Λ, a point $\mathbf{u} \in H$ and a positive real r, the discrete Gaussian probability distribution over $\Lambda + \mathbf{u}$ with parameter r is defined to be $D_{\Lambda+\mathbf{u},r}(\mathbf{x}) = \rho_r(\mathbf{x})/\rho_r(\Lambda + \mathbf{u})$, where $\rho_r(\Lambda + \mathbf{u})$ denotes $\sum_{\mathbf{y} \in \Lambda + \mathbf{u}} \rho_r(\mathbf{y})$ and more generally $\rho_r(S) = \sum_{\mathbf{y} \in S} \rho_r(\mathbf{y})$ for a countable subset $S \subset H$.

For a lattice Λ and a positive real ε, the smoothing parameter $\eta_\varepsilon(\Lambda)$ is defined to be the smallest r such that $\rho_{1/r}(\Lambda^* \setminus \{\mathbf{0}\}) \leq \varepsilon$. Given a lattice $\Lambda = \mathcal{L}(B)$, it can be shown that for any $r \geq \eta_\varepsilon(\Lambda)$, the statistical distance between the uniform distribution on the fundamental parallelepiped $\mathcal{P}(B)$ and the distribution obtained by sampling from D_r and reducing the result modulo the lattice to an element of $\mathcal{P}(B)$ is at most $\varepsilon/2$ (see Lemma 5 of [30]). Thus, if one chooses a very small value for ε, then the Gaussian distribution D_r considered over $\mathcal{P}(B)$ with $r \geq \eta_\varepsilon(\Lambda)$ behaves essentially like the uniform distribution. As shown in Claim 2.13 of [31], a lower bound for $\eta_\varepsilon(\Lambda)$ in terms of $\lambda_1(\Lambda^*)$ can be obtained by setting ε equal to the term in the sum $\sum_{\mathbf{x} \in \Lambda^* \setminus \{\mathbf{0}\}} \exp(-\pi(\eta \|\mathbf{x}\|)^2)$ corresponding to a shortest vector $\mathbf{x}$ in Λ^*. That is, $\eta_\varepsilon(\Lambda) \geq \sqrt{(-\ln \varepsilon)/\pi}/\lambda_1(\Lambda^*)$.

Remark 1. *Along with spherically symmetric distributions D_r, distributions were considered in [21] in which $\|\mathbf{x}\|^2/r^2$ is replaced by $\sum_{i=1}^{n} a_i^2/r_i^2$, where $\mathbf{x} = a_i \mathbf{h}_i$. In this paper we only use spherically symmetric distributions, which suffice for our purposes.*

A number field $K = \mathbb{Q}(\zeta)$ is obtained by adjoining ζ to $\mathbb{Q}$, where ζ is a root of a monic irreducible polynomial $f(x) \in \mathbb{Q}[x]$. The degree of $f(x)$ is the degree of the number field K. Let n denote the degree. Then K is an n-dimensional vector space over $\mathbb{Q}$. Over $\mathbb{C}$, $f(x)$ has n roots. Recall that the complex roots come in pairs and let s_1 be the number of real roots and $2s_2$ be the number of complex roots so that $s_1 + 2s_2 = n$. Suppose that the roots are ordered as $\zeta_1, \ldots, \zeta_n$, where $\zeta_1, \ldots, \zeta_{s_1}$ are real and $\zeta_{s_1+s_2+j} = \overline{\zeta_{s_1+j}}$ for $j = 1, \ldots, s_2$. Let $\sigma_i : K \to \mathbb{C}$ be the embedding of K in $\mathbb{C}$ obtained by extending the map $\zeta \mapsto \zeta_i$. The canonical embedding $\sigma : K \to \mathbb{R}^{s_1} \times \mathbb{C}^{2s_2}$ is given by $\sigma(x) = (\sigma_1(x), \ldots, \sigma_n(x))$. Note that for any $x \in K$, and $i = 1, \ldots, s_2$, $\sigma_{s_1+s_2+i}(x) = \overline{\sigma_{s_1+i}(x)}$, so that $\sigma(K) \subset H$. For $x \in K$, the trace and norm of x are respectively defined as $\mathrm{Tr}(x) = \sum_{i=1}^{n} \sigma_i(x)$ and $\mathrm{N}(x) = \prod_{i=1}^{n} \sigma_i(x)$. A geometric norm on an element $x \in K$ is defined via the embedding σ to be $\|x\| = \|\sigma(x)\|$ (and $\|x\|_\infty = \|\sigma(x)\|_\infty$).

An algebraic integer is a root of a monic polynomial with integer coefficients. Let $\mathcal{O}_K$ denote the set of all algebraic integers in the number field K. Under usual addition and multiplication in K, $\mathcal{O}_K$ forms a ring called the ring of integers of K. Any ideal (also called integral ideal) of $\mathcal{O}_K$ is also a free $\mathbb{Z}$-module of rank n, i.e., it is generated as the set of all $\mathbb{Z}$-linear combinations of some basis $\{u_1, \ldots, u_n\} \subset \mathcal{O}_K$. The norm of an ideal $\mathcal{I}$ is defined to be $\mathrm{N}(\mathcal{I}) = \#(\mathcal{O}_K/\mathcal{I})$. A fractional ideal $\mathcal{I} \subset K$ is a set such that $d\mathcal{I}$ is an integral ideal of $\mathcal{O}_K$ for some $d \in \mathcal{O}_K$. The norm of a fractional ideal $\mathcal{I}$ is defined to be $\mathrm{N}(\mathcal{I}) = \mathrm{N}(d\mathcal{I})/|\mathrm{N}(d)|$. The set of fractional ideals form a group under multiplication.

A fractional ideal $\mathcal{I}$ has a $\mathbb{Z}$-basis $\{u_1, \ldots, u_n\}$. Under the canonical embedding $\sigma(\mathcal{I})$ is a lattice, called an ideal lattice having basis $\{\sigma(u_1), \ldots, \sigma(u_n)\}$. The fractional ideal $\mathcal{I}$ is identified with its embedding $\sigma(\mathcal{I})$ and one talks of the minimum distance $\lambda_1(\mathcal{I})$ of $\mathcal{I}$ and similarly for other lattice quantities. Likewise, given $r > 0$, $D_{\mathcal{I},r}$ denotes the distribution $D_{\sigma(\mathcal{I}),r}$ over $\sigma(\mathcal{I})$. The (absolute) discriminant Δ_K of K is defined to be the square of the fundamental volume of the ideal lattice $\sigma(\mathcal{O}_K)$. The fundamental volume of any ideal lattice $\sigma(\mathcal{I})$ is $\mathrm{N}(\mathcal{I}) \cdot \sqrt{\Delta_K}$.

A lattice in K is the $\mathbb{Z}$-span of a $\mathbb{Q}$-basis of K. Let Λ be a lattice in K. The conjugate dual of Λ is defined to be $\Lambda^\vee = \{x \in K : \mathrm{Tr}(x\Lambda) \subseteq \mathbb{Z}\}$. It follows that $\sigma(\Lambda^\vee) = \overline{\sigma(\Lambda)^*}$. Let $R = \mathcal{O}_K$ which is a lattice in K. The fractional ideal $R^\vee$ is called the codifferent. For any ideal $\mathcal{I}$, $\mathcal{I}^\vee = \mathcal{I}^{-1} \cdot R^\vee$.

Ideal SVP and SIVP: An instance of the γ-approximate shortest vector problem for K, denoted K-SVP_γ, is a fractional ideal $\mathcal{I}$ in K and it is required to find a nonzero $x \in \mathcal{I}$ such that $\|x\| \leq \gamma \cdot \lambda_1(\mathcal{I})$. The γ-approximate shortest independent vector problem in K, denoted K-SIVP_γ, requires finding n linearly independent elements in $\mathcal{I}$ all of whose norms are at most $\gamma \cdot \lambda_n(\mathcal{I})$.

Ring-LWE distribution: Recall that the intuitive meaning of an LWE sample is an approximate linear equation of the form $\sum a_i s_i = b$ where b includes an error, the s_i are unknown, and the a_i and b are given to the solver. In ring-LWE $\sum a_i s_i$ is realised more efficiently because a and s are elements of a number field, and the left hand side is simply $a \cdot s$. More precisely, let K be a number field and $R = \mathcal{O}_K$. For a fractional ideal $\mathcal{J}$ in K and an integer $q \geq 2$, let $\mathcal{J}_q$ denote the set of residue classes of $\mathcal{J}$ modulo $q\mathcal{J}$. Let $\mathbb{T} = H/\sigma(R^\vee)$ denote H modulo $\sigma(R^\vee)$. Suppose $s \in R_q^\vee$ and $a \in R_q$. There are elements $x \in R^\vee$ and $y \in R$ such that $s = x + qR^\vee$ and $a = y + qR$. Define the result of the operation $a \cdot s$ to be $xy + qR^\vee$ which is in $R_q^\vee$. One can show that the operation is well defined. Similarly, the result of the operation $(a \cdot s)/q$ is defined to be $xy/q + R^\vee$ which is in $(1/q)R^\vee$ modulo $R^\vee$. By $\sigma((a \cdot s)/q)$ we will denote the element $\sigma(xy/q) + \sigma(R^\vee)$ of $\mathbb{T}$. For $s \in R_q^\vee$ and a positive real number r, a sample from the ring-LWE distribution $A_{s,r}$ over $R_q \times \mathbb{T}$ is $(a, \sigma((a \cdot s)/q) + \mathbf{e} \bmod \sigma(R^\vee))$, where a is chosen uniformly at random from R_q and $\mathbf{e}$ is chosen from H following the distribution D_r.

Remark 2. *1. We have defined the ring-LWE distribution by transferring $(a \cdot s)/q$ to H and performing the addition with the error $\mathbf{e}$ in H modulo $\sigma(R^\vee)$. This is helpful for the theoretical analysis of the reduction. In practice, on the other hand, it is computationally more efficient to transfer $\mathbf{e}$ to an approximate element in K and perform the addition in K. We briefly describe how $\mathbf{e} \in H$ can be transferred to K. Let $(\nu_1, \ldots, \nu_n)$ be a basis of K over $\mathbb{Q}$. The requirement is to find $x_1, \ldots, x_n \in \mathbb{Q}$, such that $\sigma(\sum_{i=1}^n x_i \nu_i) = \sum_{i=1}^n x_i \sigma(\nu_i)$ is close to $\mathbf{e}$. This is done as follows. Let M be the inverse of the matrix whose (i,j)-th entry is $\sigma_j(\nu_i)$, compute $(y_1, \ldots, y_n) = \mathbf{e}M \in \mathbb{R}^n$ and then choose x_i to be a rational approximation of y_i.*

2. In [21], the second component of the ring-LWE distribution is an element of the field tensor product $K \otimes_{\mathbb{Q}} \mathbb{R}$. Since $K \otimes_{\mathbb{Q}} \mathbb{R}$ and H are isomorphic as n-dimensional vector spaces over $\mathbb{R}$, we have chosen to work with the equivalent and simpler formulation where the second component of the ring-LWE is an element of H.

Search ring-LWE: Let $\alpha > 0$ be a real number and $q \geq 2$ be an integer. The search version of the ring-LWE problem, denoted ring-$\mathsf{LWE}_{q,\leq\alpha}$, is the following. For any $s \in R_q^\vee$ and a fixed positive real number $r \leq \alpha$, given access to arbitrarily many independent samples from $A_{s,r}$, find s. Formally, a probabilistic algorithm $\mathcal{A}$ to solve ring-$\mathsf{LWE}_{q,\leq\alpha}$ has access to an oracle $W_{s,r}$, where $r \leq \alpha$ is unknown, which when queried returns an independent sample from $A_{s,r}$. $\mathcal{A}$ is allowed to adaptively query $W_{s,r}$ a number of times and at the end outputs an element $s' \in R_q^\vee$. The success probability of $\mathcal{A}$ is the probability that $s' = s$. The important parameters for $\mathcal{A}$ are its success probability, its runtime and the number of times it queries its oracle.

A necessary condition for solvability of ring-$\mathsf{LWE}_{q,\leq\alpha}$ is $\alpha < \eta_\varepsilon(R^\vee)$ for all negligible ε, as otherwise the added error makes the samples essentially uniform.

Ring DLWE (fixed width): Let $r > 0$ be a real number and $q \geq 2$ be an integer. The decision version of the ring-LWE problem, denoted ring-$\mathsf{DLWE}_{q,r}$, is the following. Let s be chosen uniformly at random from $R_q^\vee$. The task is to distinguish with non-negligible advantage between arbitrarily many independent samples from $A_{s,r}$ and the same number of samples generated independently and uniformly from $R_q \times \mathbb{T}$. Formally, let $\mathcal{D}$ be a distinguisher which takes as input a list $\mathcal{T}$ consisting of elements from $R_q \times \mathbb{T}$. For a fixed value of $s \in R_q^\vee$, let $p_{s,0}$ be the probability that $\mathcal{D}$ outputs 1 when $\mathcal{T}$ consists of independent samples from $A_{s,r}$, where the probability is over all components of the input other than s as well as the internal coin tosses of $\mathcal{D}$. Let p_1 be the probability that $\mathcal{D}$ outputs 1 when $\mathcal{T}$ consists of independent samples chosen uniformly from $R_q \times \mathbb{T}$, where the probability is over the input and the internal coin tosses of $\mathcal{D}$. For a given value of s, the advantage of the distinguisher is $|p_{s,0} - p_1|$. For $\epsilon_1, \epsilon_2 \in (0, 1]$, we say that $\mathcal{D}$ is an (ϵ_1, ϵ_2)-distinguisher if $\mathcal{D}$ has advantage at least ϵ_2 for at least a proportion ϵ_1 of the set of possible $s \in R_q^\vee$,

Remark 3. *A more general definition of ring DLWE was given in [21]. In this definition, the error distribution itself is chosen from a distribution over a family of error distributions. Theorem 5.1 of [21] was proved for such a definition, while Theorem 5.2 of [21] was proved for the case where the error distribution is fixed. Since applications use a fixed error distribution, we have defined ring DLWE with fixed error distribution.*

Notation We summarise and fix some notation for future convenience.

$\log(x)$	: logarithm of x to the base 2
$\ln(x)$	: natural logarithm of x
Λ	: a lattice in H
$\lambda_1(\Lambda)$	: minimum distance of the lattice Λ
$\lambda_n(\Lambda)$	: the least real number such that Λ has n linearly independent vectors with the length of the longest being equal to this number
$\eta_\varepsilon(\Lambda)$	: smoothing parameter for a lattice Λ
K	: underlying number field
n	: degree of the number field
$\mathcal{O}_K$, R	: ring of integers of K
$\mathcal{I}$, $\mathcal{J}$	: fractional ideals of a number field
σ	: canonical embedding of a number field into H (as defined in (3))
$\mathbb{T}$	: the set of residue classes of H modulo $\sigma(R^\vee)$
$\mathcal{I}^*$	: the dual of the lattice $\sigma(\mathcal{I})$
$\mathcal{I}^\vee$	: the conjugate dual of the lattice $\sigma(\mathcal{I})$
D_r	: Gaussian distribution on H of width r
$D_{\mathcal{I},r}$	: discrete Gaussian distribution of width r on the lattice $\sigma(\mathcal{I})$
$r, r', r_i, \xi, \xi', \xi_i$	: widths of Gaussian distributions
q	: an integer ≥ 2
$\mathcal{I}_q$	: the set of residue classes of $\mathcal{I}$ modulo $q\mathcal{I}$
$\omega(\sqrt{\ln n})$	: a fixed function which grows asymptotically faster than $\sqrt{\ln n}$
α	: an upper bound on the width of Gaussian distributions that satisfies $\alpha < \sqrt{\ln n/n}$ and $\alpha q \geq 2\omega(\sqrt{\ln n})$.

3 Reducing K-SIVP_γ to search ring-$\mathsf{LWE}_{q,\leq\alpha}$

Fix three parameters: a positive integer n which will denote the degree of the number field K; an integer $q \geq 2$ which is used to define the ring-LWE problem; and a positive real number α such that $\alpha q \geq 2 \cdot \omega(\sqrt{\ln n})$. We assume $\alpha < \sqrt{\ln n/n}$ which, as remarked in [21], holds for proposed applications. In the asymptotic setting, q and α are considered to be functions of n.

The reduction of K-SIVP$_\gamma$ to ring-LWE$_{q,\leq\alpha}$ is obtained by composing reductions involving several intermediate computational problems. Let K be a number field.

- Let Γ be a function from fractional ideals in K to $\mathbb{R}$. The discrete Gaussian sampling problem in K, denoted K-DGS$_\Gamma$, is the following. Given a fractional ideal $\mathcal{I}$ in K and $r \geq \Gamma(\mathcal{I})$, produce a sample from $D_{\mathcal{I},r}$. This means producing an element of $\mathcal{I}$ such that the probability of producing $x \in \mathcal{I}$ is given by $D_{\mathcal{I},r}(x) = \rho_r(\sigma(x))/\sum_{y\in\mathcal{I}} \rho_r(\sigma(y))$.

- Given a fractional ideal $\mathcal{I}$ in K and $\xi < \lambda_1(\mathcal{I})/(2\sqrt{2n})$, an instance of the bounded distance decoding problem K-BDD$_{\mathcal{I},\xi}$ is an element $y \in K$ such that $y = x + e$, where $x \in \mathcal{I}$ and $e = \sigma^{-1}(\mathbf{e})$ is chosen according to D_ξ; the requirement is to find x' such that $x' = x$ except with negligible probability where the probability is over $\mathbf{e}$ as well as internal coin tosses. An algorithm to solve K-BDD$_{\mathcal{I},\xi}$ will take as input the pair $(\mathcal{I}, y)$. The upper bound on ξ ensures that the solution x is unique except with negligible probability. Note that ξ is unknown to the solver.

- For a fractional ideal $\mathcal{I}$ in K and an integer $q \geq 2$, the q-BDD$_{\mathcal{I},\xi}$ problem is the following. Given an instance y of K-BDD$_{\mathcal{I},\xi}$ with solution $x \in \mathcal{I}$, find $x \bmod q\mathcal{I}$.

The K-SIVP$_\gamma$ to ring-LWE$_{q,\leq\alpha}$ reduction is obtained from the following sequence of algorithms, in which $\mathcal{A}_i$ calls $\mathcal{A}_{i+1}$ as an oracle, for $0 \leq i \leq 4$.

Algorithm $\mathcal{A}_0$: Solves K-SIVP$_\gamma$ for an appropriate value of γ (see §3.2 below). The input is a fractional ideal $\mathcal{I}$ and the output is a set of n linearly independent elements of $\mathcal{I}$ the longest of which is at most $\gamma\lambda_n(\mathcal{I})$.

Algorithm $\mathcal{A}_1$: Solves K-DGS$_\Gamma$, for an appropriate Γ (see §3.2 below). The input is a pair $(\mathcal{I}, r)$, where $\mathcal{I}$ is a fractional ideal of K and $r \geq \Gamma(\mathcal{I})$. The output is a sample from the distribution $D_{\mathcal{I},r}$.

Algorithm $\mathcal{A}_2$: This is a quantum algorithm which, given as input a fractional ideal $\mathcal{I}$ and a set of samples chosen independently from $D_{\mathcal{I},r}$, returns a sample from $D_{\mathcal{I},r'}$, where $r' \leq r/2$. The conditions on r and r' are given in (27) of Appendix A.

Algorithm $\mathcal{A}_3$: Solves K-BDD$_{\mathcal{I}^\vee,\xi}$. The input is a pair $(\mathcal{I}^\vee, y)$, where $\mathcal{I}$ is a fractional ideal of K, $y = x + e$, $x \in \mathcal{I}^\vee$, and $e = \sigma^{-1}(\mathbf{e})$ is chosen according to the distribution D_ξ. Additionally, $\mathcal{A}_3$ has access to a set of samples chosen independently from $D_{\mathcal{I},r}$. The output is an $x' \in \mathcal{I}^\vee$ such that $x' = x$ except with negligible probability. The relation between r and ξ is given in (27) of Appendix A.

Algorithm $\mathcal{A}_4$: Solves q-BDD$_{\mathcal{I}^\vee,\xi}$. The input to $\mathcal{A}_4$ is the same as that to $\mathcal{A}_3$ and the output is $x' \bmod q$ such that $x' \equiv x \bmod q$ except with negligible probability.

Algorithm $\mathcal{A}_5$: Solves ring-LWE$_{q,\leq\alpha}$. Algorithm $\mathcal{A}_5$ has access to an oracle which generates samples from the ring-LWE distribution $A_{s,r}$ defined in §2, where $r \leq \alpha$ and both s and r are unknown to $\mathcal{A}_5$. The algorithm interacts with the oracle and finally outputs s.

3.1 Reducing K-SIVP$_\gamma$ to K-DGS$_\Gamma$

For a fractional ideal $\mathcal{I}$ in K, let

$$\Gamma(\mathcal{I}) \quad = \quad \frac{\gamma \cdot \lambda_n(\mathcal{I})}{2\sqrt{n}}. \tag{4}$$

Suppose $\varepsilon \leq 1/10$ and $\gamma \geq 2\sqrt{2n}\eta_\varepsilon(\mathcal{I})/\lambda_n(\mathcal{I})$, which implies that $\Gamma(\mathcal{I}) \geq \sqrt{2}\eta_\varepsilon(\mathcal{I})$. Given an algorithm $\mathcal{A}_1$ to solve K-DGS_Γ where Γ is given by (4), it is possible to construct an algorithm $\mathcal{A}_0$ to solve K-SIVP_γ. This is shown in Lemma 3.17 of [31].

We briefly review the proof. The objective is to obtain a set of n linearly independent vectors whose longest vector has length at most $2\sqrt{n}\,\Gamma(\mathcal{I})$, which using (4) is equal to $\gamma\lambda_n(\mathcal{I})$. Algorithm $\mathcal{A}_0$ uses the LLL algorithm to obtain a set B_0 of n linearly independent vectors such that the length d_0 of the longest vector in B_0 satisfies $d_0 \leq 2^{(n-1)/2}\lambda_n(\mathcal{I})$ (see Theorem 3 of [23]). From $d_0 \leq 2^{(n-1)/2}\lambda_n(\mathcal{I}) < 2^{n/2}\lambda_n(\mathcal{I})$ and (4), we have $d_0 < 2^n\Gamma(\mathcal{I})$ for $\gamma \geq \sqrt{n}/2^{n/2-1}$, where the condition on γ holds[4] for all $n \geq 8$.

For $i = 1, \ldots, n$, let $u_i = d_0/2^{i-1}$. For each i in $\{1, \ldots, n\}$, $\mathcal{A}_0$ does the following. It invokes $\mathcal{A}_1$ a total of n^2 times on the input $(\mathcal{I}, u_i)$ to obtain a set T_i of n^2 elements of $\mathcal{I}$ chosen independently from the distribution[5] $D_{\mathcal{I},u_i}$. $\mathcal{A}_0$ looks for a set B_i of n linearly independent elements in T_i. If B_i is found, then let d_i be the length of the longest vector in B_i. Finally $\mathcal{A}_0$ returns a set B_k such that d_k is the minimum of all the d_i such that B_i is defined.

The claim is that with high probability $\mathcal{A}_0$ returns a set of n linearly independent vectors whose longest vector is at most $2\sqrt{n}\,\Gamma(\mathcal{I})$. Since $d_0 < 2^n\Gamma(\mathcal{I})$, it follows that either $d_0 < \Gamma(\mathcal{I})$, or there is some $k_0 \in \{1, \ldots, n\}$ such that $\Gamma(\mathcal{I}) \leq u_{k_0} < 2\Gamma(\mathcal{I})$. If $d_0 < \Gamma(\mathcal{I})$, then $d_k \leq d_0 < \Gamma(\mathcal{I})$ and the claim holds. Otherwise, consider the k_0 such that $\Gamma(\mathcal{I}) \leq u_{k_0} < 2\Gamma(\mathcal{I})$. The conditions $\varepsilon \leq 1/10$ and $u_{k_0} \geq \Gamma(\mathcal{I}) \geq \sqrt{2}\eta_\varepsilon(\mathcal{I})$ ensure that with high probability the set T_{k_0} of n^2 vectors contains a set B_{k_0} of n linearly independent vectors (Corollary 3.16 of [31]). Further, with high probability the vectors in B_{k_0} are of length at most $u_{k_0}\sqrt{n}$ (Lemma 2.5 of [31]) which is less than $2\sqrt{n}\,\Gamma(\mathcal{I})$, i.e., $d_{k_0} < 2\sqrt{n}\,\Gamma(\mathcal{I})$. Since $d_k \leq d_{k_0}$, the claim also holds in this case. We record the following.

Proposition 1. *$\mathcal{A}_0$ invokes $\mathcal{A}_1$ a total of n^3 times.*

3.2 Reducing K-DGS_Γ to ring-$\mathsf{LWE}_{q,\leq\alpha}$

For a fractional ideal $\mathcal{I}$ in K, let

$$\gamma \;=\; 2\sqrt{2} \cdot \frac{\sqrt{n}}{\alpha} \cdot \omega(\sqrt{\ln n}) \cdot \frac{\eta_\varepsilon(\mathcal{I})}{\lambda_n(\mathcal{I})} \tag{5}$$

Since $\alpha < \sqrt{\ln n/n}$ and $\omega(\sqrt{\ln n})$ grows faster than $\sqrt{\ln n}$, the value of γ given by (5) satisfies the lower bound on γ assumed in Section 3.1. Substituting the value of γ given by (5) in (4), we obtain

$$\Gamma(\mathcal{I}) \;=\; \frac{\sqrt{2} \cdot \omega(\sqrt{\ln n}) \cdot \eta_\varepsilon(\mathcal{I})}{\alpha}. \tag{6}$$

With Γ given by (6), there is a reduction from K-DGS_Γ to ring-$\mathsf{LWE}_{q,\leq\alpha}$, where $\alpha q \geq 2\omega(\sqrt{\ln n})$. The reduction will be described in this section.

For γ given by (5), using Lemma 2.2[6] of [21] we have

$$\gamma \;\leq\; 2\sqrt{2} \cdot \frac{\sqrt{n}}{\alpha} \cdot \omega(\sqrt{\ln n})\sqrt{\ln(n/\varepsilon)}. \tag{7}$$

The relation (7) has been summarised in [21] as $\gamma = \tilde{O}(\sqrt{n}/\alpha)$.

[4] From (4), the assumptions $\Gamma(\mathcal{I}) \geq \sqrt{2}\eta_\varepsilon(\mathcal{I})$ and $\varepsilon \leq 1/10$ and the lower bound on $\eta_\epsilon(\mathcal{I})$ given in Claim 2.13 of [31], it follows that $\gamma \geq \sqrt{8\ln 10/(n\pi)} \geq \sqrt{n}/2^{n/2-1}$ for $n \geq 8$.

[5] If $d_0 < \Gamma(\mathcal{I})$, let $j = 0$, otherwise, let $j \in \{1, \ldots, n\}$ be such that $2^{j-1}\Gamma(\mathcal{I}) \leq d_0 < 2^j\Gamma(\mathcal{I})$. Then for $j+1 \leq i \leq n$, we have $u_i < \Gamma(\mathcal{I})$ implying that the pair $(\mathcal{I}, u_i)$ is an invalid input to $\mathcal{A}_1$. Since the expression for $\Gamma(\mathcal{I})$ given by (4) involves $\lambda_n(\mathcal{I})$, the value of $\Gamma(\mathcal{I})$ cannot be efficiently computed and so it is not possible to check the condition $u_i < \Gamma(\mathcal{I})$. We redefine $\mathcal{A}_1$ so that if $u_i < \Gamma(\mathcal{I})$, it returns some element of $\mathcal{I}$, but with nothing assumed about whether the selection adheres to any Gaussian distribution.

[6] Lemma 2.2 of [21] shows that $\eta_\varepsilon(\mathcal{I}) \leq \sqrt{\ln(n/\varepsilon)}\lambda_n(\mathcal{I})$.

$\mathcal{A}_1$ solves K-DGS_Γ • input $(\mathcal{I}, r)$, $r \geq \Gamma(\mathcal{I})$ • output a "sample" from $D_{\mathcal{I},r}$	$\mathcal{A}_2$ • input $\mathcal{S}$ from $D_{\mathcal{I},r}$ • output a "sample" from $D_{\mathcal{I},r'}$ • $r' = r \cdot \omega(\sqrt{\ln n})/(\alpha q) < r/2$.
$\mathcal{A}_3$ solves K-$\mathsf{BDD}_{\mathcal{I}^\vee,\xi}$ • input y such that $y = x + e$, $x \in \mathcal{I}^\vee$, $e = \sigma^{-1}(\mathbf{e})$, $\mathbf{e}$ from D_ξ • has access to $\mathcal{S}$ from $D_{\mathcal{I},r}$ • output $x' \in \mathcal{I}^\vee$ where $x' = x$ almost always • $\xi = \alpha q/(2r\omega(\sqrt{\ln n}))$	$\mathcal{A}_4$ solves q-$\mathsf{BDD}_{\mathcal{I}^\vee,\xi}$ • input same as $\mathcal{A}_3$ • outputs $x' \bmod q$
	$\mathcal{A}_5$ solves ring-$\mathsf{LWE}_{q,\leq\alpha}$ • has access to $A_{s,r}$-oracle • $r \leq \alpha$ and s are unknown to $\mathcal{A}_5$ • output s

Figure 1: Inputs and outputs of algorithms $\mathcal{A}_1$ to $\mathcal{A}_5$.

Remark 4. *As ε decreases, the expression in (7) increases. If $\varepsilon = e^{-n}$, the order of magnitude of γ is $\tilde{O}(n/\alpha)$ and not $\tilde{O}(\sqrt{n}/\alpha)$. The expression $\gamma = \tilde{O}(\sqrt{n}/\alpha)$ in [21] implies that ε is assumed not to be too small, although it has to be a negligible function of n. More precisely, suppose there is some constant d such that*

$$\varepsilon \geq n^{1-(\log n)^d}. \tag{8}$$

Then $\gamma = \tilde{O}(\sqrt{n}/\alpha)$ is justified.

We now consider the reduction of K-DGS_Γ to ring-$\mathsf{LWE}_{q,\leq\alpha}$ in further detail. While going through the description below, it will be helpful to keep in mind the inputs and outputs of the various algorithms which are shown in Figure 1.

The input to $\mathcal{A}_1$ is a pair $(\mathcal{I}, r)$, where $\mathcal{I}$ is an ideal and $r \geq \Gamma(\mathcal{I})$. Let $i_0 = 2n + \lceil (\log n)/2 \rceil$. For $i = 0, \ldots, i_0$, define $r_i = r \cdot (\alpha q/\omega(\sqrt{\ln n}))^i$. Note that $r_i \geq 2^i r$ because, by the above assumption, $\alpha q \geq 2\omega(\sqrt{\ln n})$. Also, $r_{i_0} \geq 2^{2n}\sqrt{n}r$. The lower bound on r_{i_0} ensures that $\mathcal{A}_1$ can sample from $D_{\mathcal{I},r_{i_0}}$ without requiring the help of the LWE oracle (see Lemma 3.2 of [31]). For $i = 1, \ldots, i_0$, define $\xi_i = (\alpha q)/(2r_i \cdot \omega(\sqrt{\ln n}))$.

First, $\mathcal{A}_1$ prepares a list of N samples $\mathcal{S}_{i_0}$ from $D_{\mathcal{I},r_{i_0}}$. Then for each i starting from i_0 down to 1, $\mathcal{A}_1$ invokes $\mathcal{A}_2$ a total of N times, providing it with $\mathcal{S}_i$, where each call to $\mathcal{A}_2$ returns a sample from $D_{\mathcal{I},r_{i-1}}$. To obtain a sample from $D_{\mathcal{I},r_{i-1}}$, $\mathcal{A}_2$ creates a quantum circuit and a quantum state that will produce the desired sample provided it can "uncompute" a nearest vector that is in the first register.

To accomplish the erasure of the first entangled register, we need a circuit of gates that reverse the gates in the circuit for $\mathcal{A}_3$. All of the algorithms after $\mathcal{A}_2$ will have to be incorporated into a quantum circuit that must then be reversed so that the resulting quantum circuit can be included in the quantum part of $\mathcal{A}_2$. This raises feasibility issues that we will discuss in §6.

In $\mathcal{A}_3$ the offset for the BDD instance is sampled from the distribution D_{ξ_i}. $\mathcal{A}_3$ solves the BDD instance (via $\mathcal{A}_4$) by invoking $\mathcal{A}_5$. For this, $\mathcal{A}_4$ needs to be able to simulate the responses to the LWE queries made by $\mathcal{A}_5$.

The ℓ_∞ distance of a sample from D_{ξ_i} has length at most $\xi_i' = \alpha q/(\sqrt{2}r_i)$ except with negligible probability and this is required in Lemma 4.7 of [21] to show that the simulation of responses to the LWE queries by $\mathcal{A}_4$ is correct. Below we provide further details of algorithms $\mathcal{A}_3$ and $\mathcal{A}_4$.

Algorithm $\mathcal{A}_3$ takes as input an ideal $\mathcal{I}^\vee$ and an element $y \in K$ such that $y = x + e$, where $x \in \mathcal{I}^\vee$ and $e = \sigma^{-1}(\mathbf{e})$ is chosen according to D_ξ. Additionally, it has access to a set of independent samples from $D_{\mathcal{I},r}$. Algorithm $\mathcal{A}_3$ returns x. We provide a brief overview of the construction of $\mathcal{A}_3$ using $\mathcal{A}_4$ as an oracle. Let B be a matrix whose columns form a basis for the lattice $\sigma(\mathcal{I}^\vee)$. Recall that the only difference between $\mathcal{A}_3$ and $\mathcal{A}_4$ is that the latter only finds the nearest lattice vector modulo q.

We first apply $\mathcal{A}_4(\mathcal{I}^\vee, \mathcal{S}, y)$ to find an integer vector b_1 such that $b_1 \equiv a_1 \bmod q$, where a_1 (which we do not know yet) is the integer vector such that the lattice element Ba_1 is the closest vector in $\sigma(\mathcal{I}^\vee)$ to σy. Set $y_1 = y$. Now repeat the procedure with $y_2 = (y_1 - Bb_1)/q$, obtaining $b_2 \equiv a_2 \bmod q$ such that Ba_2 is the closest vector in $\sigma(\mathcal{I}^\vee)$ to $\sigma(y_2)$. Since $B(a_1 - b_1)$ is the closest vector to $y_1 - Bb_1 = qy_2$, it follows that $a_2 = (a_1 - b_1)/q$, and that the distance between Ba_2 and y_2 is less than the distance between a_1 and b_1 by a factor of q. Applying $\mathcal{A}_4$ $n-1$ times – generating the sequences y_i and b_i, $i = 1, \ldots, n-1$ – we finally get y_n close enough to the lattice that we can use Babai's nearest plane algorithm to find the nearest lattice vector to y_n, which is a_n. Once we know a_n, as well as the $b_1, \ldots, b_n$, we can successively compute $a_i = b_i + qa_{i+1}$, $i = n-1, \ldots, 1$. Then Ba_1 is the desired output of $\mathcal{A}_3$.

The input to algorithm $\mathcal{A}_4$ is the same as that to algorithm $\mathcal{A}_3$. $\mathcal{A}_4$ returns $x \bmod q\mathcal{I}$. The general task of $\mathcal{A}_4$ in solving an instance of q-$\mathsf{BDD}_{\mathcal{I}^\vee,\xi}$ is similar to that of $\mathcal{A}_5$ in solving ring-$\mathsf{LWE}_{q,\leq\alpha}$. But there are two major differences. First, $\mathcal{A}_4$ works with the lattice $\mathcal{I}^\vee$ modulo q, whereas $\mathcal{A}_5$ works with the lattice $R^\vee$ modulo q. The reduction handles this by using an element $t \in \mathcal{I}$ that gives an isomorphism from R mod q to $\mathcal{I}$ mod q and also in the other direction between the dual lattices $\mathcal{I}^\vee$ mod q and $R^\vee$ mod q. The second difference is that $\mathcal{A}_4$ gets just one input vector $y = x + e$, whereas $\mathcal{A}_5$ has access to an oracle that provides N samples from $A_{s,r}$. Since $\mathcal{A}_4$ calls $\mathcal{A}_5$, the oracle queries made by $\mathcal{A}_5$ has to be simulated by $\mathcal{A}_4$. This is done by randomising the error in y, that is, by adding errors $\mathbf{e}'$ chosen according to $D_{\alpha/\sqrt{2}}$. For details, see §4.2 of [21].

Based on the overview and the above descriptions of algorithms $\mathcal{A}_1, \mathcal{A}_3$ and $\mathcal{A}_4$, we record the following.

Proposition 2. *1. $\mathcal{A}_1$ invokes $\mathcal{A}_2$ a total of $(2n + \lceil(\log n)/2\rceil)N$ times, where N is the number of LWE samples required by $\mathcal{A}_5$.*

2. $\mathcal{A}_2$ invokes the reverse circuit of $\mathcal{A}_3$ once.

3. $\mathcal{A}_3$ invokes $\mathcal{A}_4$ a total of n times.

4. $\mathcal{A}_4$ invokes $\mathcal{A}_5$ once.

The reduction of K-DGS_Γ to ring-$\mathsf{LWE}_{q,\leq\alpha}$ in [21] is based on the reduction of the DGS problem to the search LWE problem in [31]. There are, however, some important differences. A number of these differences pertain to the algebraic techniques needed to handle ideal lattices in [21] that do not apply to the general lattices considered in [31].

One such difference is in the distribution of the error in an LWE sample. In [31] the error follows a fixed width Gaussian distribution, while in [21] the error follows a distribution drawn from a family of elliptical Gaussian distributions. Elliptical, rather than spherical, distributions are needed in [21] to argue for the correctness of the distribution of the error in the simulated LWE samples arising in the q-$\mathsf{BDD}_{\mathcal{I}^\vee,\xi}$ to ring-$\mathsf{LWE}_{q,\leq\alpha}$ reduction. It appears that the use of elliptical distributions is a mathematical artifact introduced for the sake of the reduction rather than being of any practical importance.

3.3 The tightness gap in the K-SIVP_γ to ring-$\mathsf{LWE}_{q,\leq\alpha}$ reduction

The following theorem is a concrete version of the reduction of K-SIVP_γ to ring-$\mathsf{LWE}_{q,\leq\alpha}$ that is described in Theorem 4.1 of [21] and in the discussion following that theorem. The tightness gap in (9) follows from Propositions 1 and 2 above.

Theorem 3. *Let K be an arbitrary number field of degree n, $q \geq 2$ be a positive integer and α be a positive real number such that $\alpha q \geq 2\omega(\sqrt{\ln n})$ and $\alpha < \sqrt{\ln n/n}$. There is a quantum reduction using approximately $3n^2$ logical qubits from K-DGS_Γ, where Γ is given by (6) with $\varepsilon \leq e^{-\pi}$, to ring-$\mathsf{LWE}_{q,\leq\alpha}$. Additionally, suppose there is a positive constant d, such that $\varepsilon \geq n^{1-(\log n)^d}$. Then there is a quantum reduction from K-$\mathsf{SIVP}_{\tilde{O}(\sqrt{n}/\alpha)}$ to ring-$\mathsf{LWE}_{q,\leq\alpha}$.*

An algorithm $\mathcal{A}_0$ to solve K-$\mathsf{SIVP}_{\tilde{O}(\sqrt{n}/\alpha)}$ can be constructed using an algorithm $\mathcal{A}_5$ to solve ring-$\mathsf{LWE}_{q,\leq\alpha}$ and the number of times $\mathcal{A}_0$ calls $\mathcal{A}_5$ is approximately

$$(2n + (\log n)/2)n^4 \cdot N, \tag{9}$$

where N is the number of ring-LWE samples required by $\mathcal{A}_5$.

Remark 5. *In arriving at the expression in (9), we have assumed that the time taken by a reverse circuit for $\mathcal{A}_3$ is the same as the time taken by a circuit for $\mathcal{A}_3$.*

The term N in (9) is the number of samples that is required in ring-LWE. In practice, N would depend on α. If α is very small, then the error in the ring-LWE distribution is also very small leading to a relatively easy instance of ring-LWE. For an easy instance, obtaining about n samples might be sufficient to solve the ring-LWE problem. With a larger α, the error in the ring-LWE distribution would also be larger, which suggests that the number of samples required to solve the ring-LWE problem would be larger. Consequently, the number of oracle queries (and hence the tightness gap) grows with the difficulty of the ring-LWE problem. More concretely, if we assume $N = n^c$, then c increases as α increases. If α is small enough so that ring-LWE is easy and the tightness gap is less, then K-SIVP_γ has a larger approximation factor γ, and so becomes easier. On the other hand, as α increases, causing ring-$\mathsf{LWE}_{q,\leq\alpha}$ intuitively to become harder (the case that is relevant for cryptography), then the K-SIVP_γ problem that reduces to it becomes harder (since γ decreases), but its connection to ring-$\mathsf{LWE}_{q,\leq\alpha}$ becomes weaker because of the larger tightness gap as c increases. Because of this loosening of the connection between the two problems, it would be consistent with the reduction for the ring-$\mathsf{LWE}_{q,\leq\alpha}$ search problem to become harder as α increases but at a slower rate than the K-$\mathsf{SIVP}_{\tilde{O}(\sqrt{n}/\alpha)}$ problem. This once again shows that what the reduction gives us is somewhat less than it might seem at first glance.

4 Reducing search ring-LWE to ring-DLWE

The reduction makes use of several algebraic properties of number fields that are satisfied, in particular, by cyclotomic number fields. Before getting into the reduction, we briefly mention the relevant algebraic properties of cyclotomic number fields.

Let K be a number field and τ be an automorphism of K. One may consider τ to act on $\sigma(K)$ as follows: for $a \in K$, $\tau(\sigma(a)) = \sigma(\tau(a))$. It is possible to extend the action of τ to the whole of H in the following manner. Since τ is the identity map on $\mathbb{Q}$, it follows that τ is a linear transformation of K (considered as a vector space over $\mathbb{Q}$) to itself. If we fix a $\mathbb{Q}$-basis $\{\nu_1, \ldots, \nu_n\}$ of K, then the action of τ is given by an $n \times n$ non-singular matrix T whose entries are from $\mathbb{Q}$. More specifically, for $a \in K$, suppose $a = a_1\nu_1 + \cdots + a_n\nu_n$, with $a_i \in \mathbb{Q}$ for $i = 1, \ldots, n$. Then $\tau(a) = b_1\nu_1 + \cdots + b_n\nu_n$, where $(b_1, \ldots, b_n)^\top = T(a_1, \ldots, a_n)^\top$. Extension of τ to H is done using the matrix T. Note that $\{\sigma(\nu_1), \ldots, \sigma(\nu_n)\}$ is an $\mathbb{R}$-basis of H so that any $\mathbf{x} \in H$ can be written as $\mathbf{x} = x_1\sigma(\nu_1) + \cdots + x_n\sigma(\nu_n)$, where $x_i \in \mathbb{R}$ for $i = 1, \ldots, n$. Then $\tau(\mathbf{x})$ is defined to be $\mathbf{y} = y_1\sigma(\nu_1) + \cdots + y_n\sigma(\nu_n)$, where $(y_1, \ldots, y_n)^\top = T(x_1, \ldots, x_n)^\top$. Extending an automorphism τ of K to H allows us to apply τ to samples drawn from an error distribution defined over H.

Cyclotomic number fields. For $m \geq 1$, let $\Phi_m(x)$ be the m-th cyclotomic polynomial having degree $n = \varphi(m)$. The m-th cyclotomic number field K is $\mathbb{Q}(\zeta)$, where ζ is a root of $\Phi_m(x)$. Let $R = \mathcal{O}_K$. Henceforth, only cyclotomic number fields will be considered.

The power basis $\{1, \zeta, \zeta^2, \ldots, \zeta^{n-1}\}$ for K over $\mathbb{Q}$ is also a $\mathbb{Z}$-basis for $\mathcal{O}_K = \mathbb{Z}[x]/\Phi(x)$. Let $\mathcal{I}$ be a fractional ideal in $\mathcal{O}_K$ and v be a shortest nonzero element in $\mathcal{I}$. Multiplying v by $1, \zeta, \ldots, \zeta^{n-1}$ gives a set of n linearly independent vectors of the same length, and hence $\lambda_n(\mathcal{I}) = \lambda_1(\mathcal{I})$. Consequently, a solution to K-SIVP_γ provides a solution to K-SVP_γ and vice versa. This may be contrasted with the case for general lattices, where it has been shown in [22] that $\mathsf{SIVP}_{\sqrt{n}\gamma}$ reduces in polynomial time to SVP_γ, but not that SIVP_γ reduces to SVP_γ.

Let q be a prime number such that $q = 1 \bmod m$ so that $q = km + 1$ for some non-negative integer k. Noting that $\mathbb{Z}_q^* = \langle g \rangle$ for a generator g, it follows that the element $\omega = g^k$ has order m in $\mathbb{Z}_q^*$. The m-th cyclotomic polynomial factors over $\mathbb{Z}_q$ as $\Phi_m(x) = \prod_{i \in \mathbb{Z}_m^*}(x - \omega^i)$. Consequently, $\langle q \rangle = \prod_{i \in \mathbb{Z}_m^*} \mathfrak{q}_i$, where $\mathfrak{q}_i = \langle q, x - \omega^i \rangle$ is a prime ideal of $\mathcal{O}_K$ having norm q. (Note that the ideals $\mathfrak{q}_i$ have been indexed by elements of $\mathbb{Z}_m^*$ rather than by the integers $\{1, \ldots, n\}$.)

The field K has n automorphisms $\tau_k(\zeta) = \zeta^k$, for $k \in \mathbb{Z}_m^*$. It follows[7] that for $k \in \mathbb{Z}_m^*$, $\tau_k(\mathfrak{q}_i) = \mathfrak{q}_{ik^{-1} \bmod m}$ and $\tau_k^{-1} = \tau_{k^{-1} \bmod m}$. Also, for $k \in \mathbb{Z}_m^*$, R and $R^\vee$ are fixed by τ_k and so $\tau_k(R_q) = R_q$. Hence, if a is distributed uniformly in R_q, then $\tau_k(a)$ is also distributed uniformly in R_q.

For $i \in \mathbb{Z}_m^*$, it can be shown that the quotient group $R^\vee/(\mathfrak{q}_i R^\vee)$ has cardinality q and the representatives of the q distinct cosets can be taken to be the elements of the set $\{0, \ldots, q-1\}$. The cardinality of the set $R_q^\vee$ is q^n. Using the Chinese Remainder Theorem (CRT), it can be shown that there is an isomorphism $\mathfrak{I}$ from $R_q^\vee$ to $\oplus_{i \in \mathbb{Z}_m^*}(R^\vee/(\mathfrak{q}_i R^\vee))$. Further, $\mathfrak{I}$ can be efficiently computed in both the forward and the backward directions. For $i \in \mathbb{Z}_m^*$, let $w_i \in \{0, \ldots, q-1\}$ represent a coset of $R^\vee/(\mathfrak{q}_i R^\vee)$. Given $(w_i)_{i \in \mathbb{Z}_m^*}$, it is possible to efficiently construct $w \in R_q^\vee$ such that the i-th component of $\mathfrak{I}(w)$ is represented by w_i. For the sake of notational convenience, we let w denote $\mathfrak{I}^{-1}((w_i)_{i \in \mathbb{Z}_m^*})$.

4.1 Intermediate problems

The "search to decision" reduction is obtained by composing several individual reductions between intermediate problems. The search ring-LWE problem requires finding s in $R_q^\vee$. The first step of the reduction is to show that it is sufficient to find any one of the components in the image of s under the isomorphism $\mathfrak{I}$. The relevant intermediate problem is the following.

Ring-LWE over $\mathfrak{q}_i$. For $i \in \mathbb{Z}_m^*$, the $\mathfrak{q}_i$-LWE$_{q,\leq\alpha}$ problem is the following. For $s \in R_q^\vee$ and a positive real number $r \leq \alpha$, given access to samples from $A_{s,r}$, the requirement is to find the i-th component of $\mathfrak{I}(s)$.

Let the representatives of $\mathbb{Z}_m^*$ be chosen from the set $\{1, \ldots, m-1\}$ with the usual ordering. For $i \in \mathbb{Z}_m^*$, let $i-$ denote the largest element in $\mathbb{Z}_m^*$ which is less than i with the convention that $1-$ is taken to be 0.

The distribution $A_{s,r}^i$. For $i \in \mathbb{Z}_m^* \cup \{0\}$, $s \in R_q^\vee$ and a positive real number r, the distribution $A_{s,r}^i$ over $R_q \times \mathbb{T}$ is defined in the following manner. A sample from $A_{s,r}$ consists of a pair $(a, \mathbf{b})$, where $a \in R_q$ and $\mathbf{b} = \sigma((a \cdot s)/q) + \mathbf{e} \bmod \sigma(R^\vee)$. A sample from $A_{s,r}^i$ is a sample from $A_{s,r}$ whose k-th component for $k \leq i$ has been randomised by adding a uniform random $h_k \in \{0, 1, \ldots, q-1\}$ to the k-th component of $a \cdot s$, thereby hiding the information[8].

A sample from $A_{s,r}^i$ hides information about s with respect to the factors $\mathfrak{q}_k$ of $\langle q \rangle$ for $k \in \mathbb{Z}_m^*$ and $k \leq i$. For $i \in \mathbb{Z}_m^*$ or $i = 0$, as i increases from 0 to $m-1$, information about s is hidden in one more $\mathfrak{q}_i$-component than in the previous step. At the beginning, i.e. $i = 0$, all the components in the output of $\mathfrak{I}$ carry information about s, while at the end, i.e., $i = m-1$, the element $a \cdot s + h$ is a uniform random element of $R_q^\vee$ which is independent of both s and a. So for a sample $(a, \mathbf{b})$ drawn from $A_{s,r}^{m-1}$, a is uniform over R_q and $\mathbf{b}$ is independent of a; further, $\mathbf{b}$ is the sum modulo $\sigma(R^\vee)$ of a uniform random element of $\sigma(R_q^\vee)/q$ and an element drawn from the distribution D_r. Consequently, a sample drawn from $A_{s,r}^{m-1}$ is almost uniform over $R_q \times \mathbb{T}$ (see Lemma 5.13 of [21]).

[7] The following fact is used to obtain $\tau_k(\mathfrak{q}_i) = \mathfrak{q}_{ik^{-1} \bmod m}$: for j such that $i \equiv jk \bmod m$, $\tau_k(\zeta - \omega^i) = \tau_k(\zeta - \omega^{jk}) = \zeta^k - \omega^{jk} = (\zeta - \omega^j)x$, where $x = \zeta^{k-1} + \omega^j \zeta^{k-2} + \cdots + \omega^{j(k-1)}$ is in R.

[8] For $i \in \mathbb{Z}_m^*$, let $\chi(i)$ be the following distribution over $R_q^\vee$. For $k \in \mathbb{Z}_m^*$, choose $h_k \in \{0, \ldots, q-1\}$ as follows: $h_k = 0$ for $k > i$; and for $k \leq i$, the h_k's are chosen independently and uniformly; return $h = \mathfrak{I}^{-1}((h_k)_{k \in \mathbb{Z}_m^*})$.

A sample from $A_{s,r}^i$ is obtained as follows. For $i \in \mathbb{Z}_m^*$, let h be sampled from $\chi(i)$. Choose $(a, \mathbf{b}) \leftarrow A_{s,r}$ and output $(a, \mathbf{b} + \sigma(h)/q \bmod \sigma(R^\vee))$ as a sample from $A_{s,r}^i$; for $i = 0$, the distribution $A_{s,r}^0$ is defined to be $A_{s,r}$.

$\mathcal{A}_5$ solves ring-$\mathsf{LWE}_{q,\leq\alpha}$ • has access to $A_{s,r}$-oracle • $r \leq \alpha$ and s are unknown to $\mathcal{A}_5$ • output s	$\mathcal{A}_6$ solves $\mathfrak{q}_i$-$\mathsf{LWE}_{q,\leq\alpha}$, $i \in \mathbb{Z}_m^*$ • has access to $A_{s,r}$-oracle • $r \leq \alpha$ and s are unknown to $\mathcal{A}_5$ • output the i-th component of $\mathfrak{I}(s)$
$\mathcal{A}_7$ solves ring-$\mathsf{VWDLWE}^i_{q,\leq\alpha}$, $i \in \mathbb{Z}_m^*$ • has access to $A^j_{s,r}$, $s \in R_q^\vee$, $r \leq \alpha$, $j \in \{i, i-\}$ • output j	
$\mathcal{D}_1$ solves ring-$\mathsf{DLWE}^i_{q,r}$ • distinguishes between $A^i_{s,r}$ and $A^{i-}_{s,r}$ • $s \in R_q^\vee$, $r \leq \alpha$	$\mathcal{D}_2$ solves ring-$\mathsf{DLWE}_{q,r}$ • distinguishes between $A_{s,r}$ and $U(R_q \times \mathbb{T})$ • $s \in R_q^\vee$, $r \leq \alpha$

Figure 2: Inputs and outputs of algorithms $\mathcal{A}_5$ to $\mathcal{A}_7$ and distinguishers $\mathcal{D}_1$ and $\mathcal{D}_2$.

Variable width ring-DLWE relative to $\mathfrak{q}_i$. For $i \in \mathbb{Z}_m^*$ and a positive real number α, the ring-$\mathsf{VWDLWE}^i_{q,\leq\alpha}$ problem is the following. Given access to $A^j_{s,r}$ for $s \in R_q^\vee$, positive real number $r \leq \alpha$ and $j \in \{i, i-\}$, the requirement is to find j. In other words, the solver must determine whether or not the i-th component of the distribution has been randomised.

Remark 6. *The letters* VWD *before* LWE *denote 'variable width decision'. In Definition 5.8 of [21] this problem was denoted* WDLWE*, meaning worst-case decision LWE. In our view, the use of "worst-case" is inappropriate, whereas "variable width" indicates an important feature of the problem.*

The next step is to consider a fixed width version of the DLWE problem with respect to the ideal $\mathfrak{q}_i$.

Ring DLWE (fixed width) relative to $\mathfrak{q}_i$. Let $i \in \mathbb{Z}_m^*$ and $r_0 > 0$ be a real number. The ring-DLWE^i_{q,r_0} problem is the following. Choose s uniformly at random from $R_q^\vee$. The requirement is to distinguish between inputs from A^{i-}_{s,r_0} and A^i_{s,r_0}. Formally, let $\mathcal{D}_1$ be an algorithm which takes as input a list $\mathcal{T}$ of samples from $A^j_{s,r}$ with $j \in \{i-, i\}$ and outputs a bit. For a fixed $s \in R_q^\vee$, let $p_{s,0}$ (resp. $p_{s,1}$) be the probability that $\mathcal{D}_1$ outputs 1 when $\mathcal{T}$ consists of samples from A^{i-}_{s,r_0} (resp. A^i_{s,r_0}), where the probability is taken over all components of the input other than s as well as the internal coin tosses of $\mathcal{D}_1$. The advantage of the distinguisher is $|p_{s,0} - p_{s,1}|$. For $\epsilon_1, \epsilon_2 \in (0, 1]$, we say that $\mathcal{D}_1$ is an (ϵ_1, ϵ_2)-distinguisher if $\mathcal{D}_1$ has advantage at least ϵ_2 for at least a proportion ϵ_1 of the set of possible $s \in R_q^\vee$,

Remark 7. *The above definition is based on Definition 5.10 of [21]. The formulation is different from that of Definition 5.10 and is in the form that is actually used in Lemma 5.16 and Theorem 3.6 of [21].*

The search ring-LWE to ring-DLWE reduction involves the following algorithms.

$\mathcal{A}_6$: an algorithm to solve $\mathfrak{q}_i$-$\mathsf{LWE}_{q,\leq\alpha}$
$\mathcal{A}_7$: an algorithm to solve ring-$\mathsf{VWDLWE}^i_{q,\leq\alpha}$
$\mathcal{D}_1$: a distinguisher for ring-$\mathsf{DLWE}^i_{q,r}$
$\mathcal{D}_2$: a distinguisher for ring-$\mathsf{DLWE}_{q,r}$

While going through the description below, it will be helpful to keep in mind the inputs and outputs of the various algorithms which are shown in Figure 2.

$\mathcal{A}_6$ has access to an oracle which returns samples from $A_{s,r}$ for some unknown s and unknown $r \leq \alpha$; $\mathcal{A}_7$ has access to an oracle $A^j_{s,r}$ for some unknown s, unknown $r \leq \alpha$ and j equal to either i or $i-$.

The overall reduction proceeds in several steps to construct an algorithm $\mathcal{A}_5$ to solve ring-LWE by using a distinguisher $\mathcal{D}_2$ for ring-DLWE_{q,r_0} as an oracle. In the context of cryptography, ring-DLWE_{q,r_0} is the problem whose solution breaks the cryptosystem, and $\mathcal{A}_5$, in turn, is used by $\mathcal{A}_0$ via $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$, and $\mathcal{A}_4$ to solve SIVP_γ.

In the first step, $\mathcal{A}_5$ is constructed by using an algorithm $\mathcal{A}_6$ as an oracle to solve $\mathfrak{q}_i$-$\mathsf{LWE}_{q,\leq\alpha}$ for some i. The idea is to individually compute the components of $\mathfrak{I}(s)$ using $\mathcal{A}_6$. Since $\mathcal{A}_6$ works for a particular i, the automorphisms of the number field are used to ensure that the j-th component of $\mathfrak{I}(s)$ is transferred to the i-th component so that $\mathcal{A}_6$ can be applied. In the next step, $\mathcal{A}_6$ is constructed by using an algorithm $\mathcal{A}_7$ as an oracle to solve ring-$\mathsf{VWDLWE}^i_{q,\leq\alpha}$. There are q possible x values of the i-th component. For each x, $\mathcal{A}_6$ modifies the LWE samples in a manner such that if x is the correct value of $s \bmod \mathfrak{q}_i R^\vee$, then the samples are from the distribution $A^{i-}_{s,r}$, while if x is not equal to $s \bmod \mathfrak{q}_i R^\vee$, then the samples are from the distribution $A^i_{s,r}$. The oracle $\mathcal{A}_7$ can be used to determine which of these two cases occurs.

The biggest step in the reduction is the construction of $\mathcal{A}_7$ using a distinguisher $\mathcal{D}_1$ for ring-DLWE^i_{q,r_0} as an oracle. $\mathcal{D}_1$ is an oracle that, given two ℓ-tuples of samples that are randomised in the first i (resp. $i-$) components of H and come from LWE-sampling with known Gaussian error distribution in the remaining components, can distinguish between them. The construction of $\mathcal{A}_7$ shows that using $\mathcal{D}_1$, one can answer the same question when the LWE-sampling is with a Gaussian error distribution of width that is unknown (but less than a known bound). The last step is the construction of $\mathcal{D}_1$ by using a distinguisher $\mathcal{D}_2$ for ring-DLWE_{q,r_0} as an oracle.

Further details of the various steps are provided below with a focus on concrete aspects.

4.2 Reducing ring-$\mathsf{LWE}_{q,\leq\alpha}$ to $\mathfrak{q}_i$-$\mathsf{LWE}_{q,\leq\alpha}$

Suppose $\mathcal{A}_6$ is an algorithm to solve $\mathfrak{q}_i$-$\mathsf{LWE}_{q,\leq\alpha}$ for some particular $i \in \mathbb{Z}^*_m$. We provide a brief description of the construction of Algorithm $\mathcal{A}_5$ using $\mathcal{A}_6$ as an oracle (see Lemma 5.5 of [21] for details). The goal of $\mathcal{A}_5$ is to compute s. This can be done if each component of $\mathfrak{I}(s)$ can be computed. Algorithm $\mathcal{A}_6$ can compute the i-th component of $\mathfrak{I}(s)$. $\mathcal{A}_5$ uses automorphisms to map the j-th component to the i-th component in the following manner. For each $j \in \mathbb{Z}^*_m$, let $k = j \cdot i^{-1} \bmod m$. $\mathcal{A}_5$ then invokes $\mathcal{A}_6$, and whenever $\mathcal{A}_6$ makes a query for an LWE sample, $\mathcal{A}_5$ queries its own LWE oracle to obtain a sample $(a, \mathbf{b})$. It responds to the query by sending $(\tau_k(a), \tau_k(\mathbf{b}))$ to $\mathcal{A}_6$. Let the output of $\mathcal{A}_6$ on the j-th invocation be denoted as s_j. Finally, $\mathcal{A}_5$ returns $\mathfrak{I}^{-1}((s_j)_{j\in\mathbb{Z}^*_m})$.

Proposition 4. *$\mathcal{A}_5$ invokes $\mathcal{A}_6$ a total of n times. The numbers of LWE queries made by $\mathcal{A}_5$ and $\mathcal{A}_6$ are equal.*

It is interesting to note that the reduction of ring-$\mathsf{LWE}_{q,\leq\alpha}$ to $\mathfrak{q}_i$-$\mathsf{LWE}_{q,\leq\alpha}$ is made possible because ideal lattices for cyclotomic number fields possess some nice algebraic properties. On the other hand, the reduction itself can be considered to be a step in a possible attack on the search ring-LWE problem. This is because thanks to the automorphisms, if an algorithm evaluates the i-th component for any fixed i, then it is possible to use it to evaluate all the components; and for a fixed i the search space is not too big for exhaustive search.

The construction of $\mathcal{A}_5$ from $\mathcal{A}_6$ is based on two points, the existence of n automorphisms and the split of $\langle q\rangle$ into linear factors. There are two directions in which one could generalise from cyclotomic fields with primes q that split completely. First, in a cyclotomic field one can take any prime q that does not divide m, in which case $\langle q\rangle$ splits into n/f distinct prime ideals where f is the residue field degree. In other words, $R/(\mathfrak{q}_i R)$ is no longer the field of q elements, but rather is a degree-f extension. In that case there would still be a contribution of n to the tightness gap, because while there are only n/f prime ideals, for each prime ideal it would be required to find f coordinates, so the Chinese Remainder Theorem presumably takes as much work as before (or more). This approach has been briefly mentioned in footnote 8 on Page 26 of [21]. Second, one can generalize to non-cyclotomic Galois fields. There it is still possible to find primes q that split completely, and if q does not split completely, there will again be n/f prime ideals with residue field of degree f (provided q does not divide the discriminant of the field).

4.3 Reducing $\mathfrak{q}_i$-LWE$_{q,\leq\alpha}$ to ring-VWDLWE$^i_{q,\leq\alpha}$

Suppose $\mathcal{A}_7$ is an algorithm to solve ring-VWDLWE$^i_{q,\leq\alpha}$. This algorithm is used as an oracle to construct an algorithm $\mathcal{A}_6$ to solve $\mathfrak{q}_i$-LWE$_{q,\leq\alpha}$. We provide a brief description of the construction and for details we refer to Lemma 5.9 of [21]. The requirement for $\mathcal{A}_6$ is to determine the i-th component of $\mathfrak{I}(s)$. As mentioned earlier, each component of $\mathfrak{I}(s)$ can be represented by an element from the set $\{0,\ldots,q-1\}$. For each $x \in \{0,\ldots,q-1\}$, $\mathcal{A}_6$ does the following. It first computes an element $g \in R_q^\vee$ such that $\mathfrak{I}(g)$ is equal to x in the i-th component and is equal to zero in all other components. Then $\mathcal{A}_6$ invokes $\mathcal{A}_7$. For each LWE query made by $\mathcal{A}_7$, $\mathcal{A}_6$ queries its own oracle to obtain a pair $(a,\mathbf{b})$. It then computes an element $v \in R_q$ such that under the isomorphism from R_q to $\oplus_{j\in\mathbb{Z}_m^*} R/(\mathfrak{q}_i R)$, the i-th component is chosen uniformly at random from $\{0,\ldots,q-1\}$ and all other components are equal to zero. Next, $\mathcal{A}_6$ adds $(v, \sigma((v\cdot g)/q) \bmod \sigma(R^\vee))$ to the sample $(a,\mathbf{b})$ and then randomises the first $i-$ components by adding a random element to the second part of each of those components of the sample. $\mathcal{A}_6$ sends the resulting modified sample to $\mathcal{A}_7$ as its response to the query[9]. At the end of its oracle queries, if $\mathcal{A}_7$ returns $i-$, then $\mathcal{A}_6$ returns x. The crucial point for correctness is that if the value of x is equal to the i-th component of $\mathfrak{I}(s)$, then the samples returned to $\mathcal{A}_7$ are from the distribution $A^{i-}_{s,r}$, and if not, then the samples returned to $\mathcal{A}_7$ are from the distribution $A^i_{s,r}$.

Proposition 5. *$\mathcal{A}_6$ invokes $\mathcal{A}_7$ at most q times. The numbers of LWE queries made by $\mathcal{A}_6$ and $\mathcal{A}_7$ are equal.*

4.4 Reducing ring-VWDLWE$^i_{q,\leq\alpha}$ to ring-DLWE$^i_{q,\mathfrak{r}}$

For a fixed $i \in \mathbb{Z}_m^*$, Lemma 5.16 of [21] states[10] that ring-VWDLWE$^i_{q,\leq\alpha}$ reduces to ring-DLWE$^i_{q,\mathfrak{r}}$ in randomised polynomial time, where

$$\mathfrak{r} \;=\; \alpha\cdot\left(\frac{n\ell}{\ln(n\ell)}\right)^{1/4} \tag{10}$$

and ℓ is the number of LWE samples required by the distinguisher for ring-DLWE$^i_{q,\mathfrak{r}}$. Note that (10) shows that $\mathfrak{r} > \alpha$. Since our goal is to estimate feasibility, we perform a concrete analysis which turns out to be considerably more complicated than the sketch of a proof in the asymptotic setting that was provided in [21]. In particular, in order to get a rigorous proof for this part of the reduction we needed to include the factor N_2, defined below.This has a substantial effect on the ratio $\mathfrak{r}/\alpha$, which now becomes

$$\frac{\mathfrak{r}}{\alpha} \;=\; \left(\frac{nN_2\ell}{\ln(nN_2\ell)}\right)^{1/4}, \tag{11}$$

and that, in turn, brings a new term $N_2^{1/4} > \sqrt{n}$ (see below) into the SIVP approximation factor.

The essential difference between the problems ring-VWDLWE$^i_{q,\leq\alpha}$ and ring-DLWE$^i_{q,\mathfrak{r}}$ is in the distribution of the error of the LWE samples. For the former problem the errors follow D_r, where $r \leq \alpha$, while in the latter problem the errors follow $D_\mathfrak{r}$ with $\mathfrak{r} > \alpha$ as mentioned above. The reduction of ring-VWDLWE$^i_{q,\leq\alpha}$ to ring-DLWE$^i_{q,\mathfrak{r}}$ is a trade-off between narrower width with no knowledge of the width except for an upper bound versus wider width (which generally means less useful samples) with knowledge of the width.

Let $\mathcal{T} = ((a_k,\mathbf{b}_k))_{1\leq k\leq\ell}$ be a list of ℓ samples from $A^j_{s,r}$, where $r \leq \alpha$ and j is equal to either $i-$ or i. Suppose $t \in R_q^\vee$ and let $\mathbf{f}_1,\ldots,\mathbf{f}_\ell$ be chosen independently from $D_\mathfrak{r}$. Define a list $\mathcal{T}' = ((a'_k,\mathbf{b}'_k))_{1\leq k\leq\ell}$, where $a'_k = a_k$ and $\mathbf{b}'_k = \mathbf{b}_k + \sigma(a_k\cdot t)/q + \mathbf{f}_k \pmod{\sigma(R^\vee)}$. Since $(a_k,\mathbf{b}_k)$ is a sample from $A^j_{s,r}$, the error vector $\mathbf{e}_k$ in $\mathbf{b}_k$ follows

[9] In other words, $\mathcal{A}_6$ generates h from the distribution $\chi(i-)$ and returns the pair $(a+v, \mathbf{b}+(\sigma(h+v\cdot g))/q \bmod \sigma(R^\vee))$ to $\mathcal{A}_7$.

[10] The actual statement of Lemma 5.16 in [21] is for the case where the error distribution for the ring-VWDLE problem is from a family of elliptical Gaussian distributions. Here we consider the simpler situation where the error distribution is from a family of spherical Gaussian distributions. This simplification does not have any effect on the concrete security analysis.

D_r. The error in $\mathbf{b}'_k$, which is $\mathbf{e}_k + \mathbf{f}_k$, follows $D_{r'}$, where $r' = \sqrt{r^2 + \mathfrak{r}^2}$. Hence, the samples in $\mathcal{T}'$ are from the distribution $A^j_{s+t,r'}$.

From $\mathcal{T}'$, a list $\mathcal{T}''$ is obtained as follows. For each pair $(a'_k, \mathbf{b}'_k)$ in $\mathcal{T}'$, a'_k is unchanged and $\mathbf{b}'_k$ is modified so as to partially randomise $\mathbf{b}_k$; namely all components of of $a_k \cdot (s+t)$ up through the i-th are randomised[11]. Note that irrespective of the value of j, the samples in $\mathcal{T}''$ are from the distribution $A^i_{s+t,r'}$. Note that the LWE secret in both $\mathcal{T}'$ and $\mathcal{T}''$ is $s+t$ and the distribution of the LWE errors is $D^\ell_{r'}$. This ensures two things. The first is a random self-reduction where the LWE secret s is mapped to $s+t$ and the second is the addition of the $\mathbf{f}$-errors so as to get the error width r' to within a small multiplicative factor of $\mathfrak{r}$.

Let $\mathcal{D}_1$ be an (ϵ_1, ϵ_2)-distinguisher for ring-$\mathsf{DLWE}^i_{q,\mathfrak{r}}$. Using $\mathcal{D}_1$, an algorithm $\mathcal{A}_7$ for ring-$\mathsf{VWDLWE}^i_{q,\leq\alpha}$ is constructed as follows. $\mathcal{A}_7$ has access to an oracle for $A^j_{s,r}$, where $r \leq \alpha$ and j is either i or $i-$. The requirement is to determine j. The construction of $\mathcal{A}_7$ has two nested loops: an outer loop of N_1 iterations and for each of these iterations an inner loop of N_2 iterations. In each iteration of the outer loop, $\mathcal{A}_7$ chooses t uniformly at random from $R^\vee_q$. Then the inner loop of N_2 iterations starts. In each of the N_2 iterations of the inner loop, $\mathcal{A}_7$ obtains a list $\mathcal{T}$ of samples from $A^j_{s,r}$; chooses $\mathbf{f}_1, \ldots, \mathbf{f}_\ell$ independently from $D^\ell_{\mathfrak{r}}$; and uses $\mathcal{T}$, t and $\mathbf{f}_1, \ldots, \mathbf{f}_\ell$ to prepare the lists $\mathcal{T}'$ and $\mathcal{T}''$ as described above. Then $\mathcal{A}_7$ runs $\mathcal{D}_1$ on $\mathcal{T}'$ and $\mathcal{T}''$ obtaining in return the corresponding 1-bit outputs. At the end of the inner loop of N_2 iterations, $\mathcal{A}_7$ obtains estimates $\hat{\mathfrak{p}}_0$ and $\hat{\mathfrak{p}}_1$ of the probabilities $\mathfrak{p}_0$ and $\mathfrak{p}_1$ that $\mathcal{D}_1$ accepts inputs from the distributions $A^j_{s+t,r'}$ and $A^i_{s+t,r'}$ respectively. If $|\hat{\mathfrak{p}}_0 - \hat{\mathfrak{p}}_1| \geq \epsilon_2/4$, then $\mathcal{A}_7$ returns $i-$ and stops. If in none of the N_1 iterations the condition $|\hat{\mathfrak{p}}_0 - \hat{\mathfrak{p}}_1| \geq \epsilon_2/4$ is satisfied, then $\mathcal{A}_7$ returns i. Figure 2 in Appendix B provides a pseudo-code description of $\mathcal{A}_7$.

Note that $\mathcal{D}_1$ is supposed to work for errors following the distribution $D^\ell_{\mathfrak{r}}$. It is, however, invoked on the lists $\mathcal{T}'$ and $\mathcal{T}''$, where the errors in the samples in these lists follow $D^\ell_{r'}$. Due to the change in the width of the error distribution, $\mathcal{D}_1$ may not behave as an (ϵ_1, ϵ_2)-distinguisher. This is taken care of in the following analysis.

Algorithm $\mathcal{A}_7$ fails if it returns an incorrect answer. This can happen in two ways, namely that $j = i$ and $\mathcal{A}_7$ returns $i-$, and $j = i-$ and $\mathcal{A}_7$ returns i. We call the former to be Type-1 failure and the latter to be Type-2 failure.

Case $j = i$. In this case, the samples in both $\mathcal{T}'$ and $\mathcal{T}''$ follow $A^i_{s+t,r'}$ and so $\mathfrak{p}_0 = \mathfrak{p}_1$. Consider any one of the N_1 iterations of the outer loop. From the additive form of the Chernoff-Hoeffding bound [17], we have

$$\begin{aligned} \Pr[\mathfrak{p}_0 - \epsilon_2/8 \leq \hat{\mathfrak{p}}_0 \leq \mathfrak{p}_0 + \epsilon_2/8] &\geq 1 - 2\exp(-N_2\epsilon_2^2/32), \\ \Pr[\mathfrak{p}_1 - \epsilon_2/8 \leq \hat{\mathfrak{p}}_1 \leq \mathfrak{p}_1 + \epsilon_2/8] &\geq 1 - 2\exp(-N_2\epsilon_2^2/32). \end{aligned}$$

Since $\mathfrak{p}_0 = \mathfrak{p}_1$, it follows that $\Pr[|\hat{\mathfrak{p}}_0 - \hat{\mathfrak{p}}_1| \leq \epsilon_2/4] \geq 1 - 4\exp(-N_2\epsilon_2^2/32)$. So in any of the N_1 iterations of the outer loop the probability of Type-1 failure is at most $4\exp(-N_2\epsilon_2^2/32)$. The probability of Type-1 failure in any of the N_1 iterations is then at most $4N_1\exp(-N_2\epsilon_2^2/32)$.

Case $j = i-$. In this case, the samples in $\mathcal{T}'$ follow $A^{i-}_{s+t,r'}$ while the samples in $\mathcal{T}''$ follow $A^i_{s+t,r'}$. In any of the N_2 iterations of the inner loop, let $\mathbf{z}_1 = \mathbf{e}_1 + \mathbf{f}_1, \ldots, \mathbf{z}_\ell = \mathbf{e}_\ell + \mathbf{f}_\ell$ be the errors in the LWE samples in the lists $\mathcal{T}'$ and $\mathcal{T}''$. Let $\mathbf{z}$ be a vector consisting of all the $N_2\ell$ errors in the N_2 iterations of the loop. The vector $\mathbf{z}$ follows $D^{\ell N_2}_{r'}$. Suppose instead that it follows $D^{\ell N_2}_{\mathfrak{r}}$. Later we will compute a correction factor to account for the width being r' rather than $\mathfrak{r}$. We denote the corresponding probabilities and their estimates by p_0, p_1, $\hat{p}_0$ and $\hat{p}_1$. Let $p_{s+t,0}$ and $p_{s+t,1}$ respectively denote the probabilities p_0 and p_1 corresponding to a particular value of t. Similarly, let $\hat{p}_{s+t,0}$ and $\hat{p}_{s+t,1}$ respectively denote the estimates $\hat{p}_0$ and $\hat{p}_1$ corresponding to a particular value of t. Further, let $\hat{p}_{s+t,\mathbf{z},0}$ and $\hat{p}_{s+t,\mathbf{z},1}$ denote these estimates for a particular value of t and $\mathbf{z}$.

We say that a value $s+t$ is good if $|p_{s+t,0} - p_{s+t,1}| \geq \epsilon_2$. From the definition of an (ϵ_1, ϵ_2)-distinguisher, the probability of a good $s+t$ is at least ϵ_1. For a good $s+t$, using the additive form of the Chernoff-Hoeffding

[11] In other words, $(a'_k, \mathbf{b}'_k)$ is modified to $(a'_k, \mathbf{b}'_k + \sigma(h_k)/q \pmod{\sigma(R^\vee)})$, where h_k is chosen from $\chi(i)$.

bound [17], we have

$$\begin{aligned}\Pr[p_{s+t,0} - \epsilon_2/4 \le \hat{p}_{s+t,0} \le p_{s+t,0} + \epsilon_2/4] &\ge 1 - 2\exp(-N_2\epsilon_2^2/8),\\ \Pr[p_{s+t,1} - \epsilon_2/4 \le \hat{p}_{s+t,1} \le p_{s+t,1} + \epsilon_2/4] &\ge 1 - 2\exp(-N_2\epsilon_2^2/8).\end{aligned}$$

Since the events $p_{s+t,0} - \epsilon_2/4 \le \hat{p}_{s+t,0} \le p_{s+t,0} + \epsilon_2/4$ and $p_{s+t,1} - \epsilon_2/4 \le \hat{p}_{s+t,1} \le p_{s+t,1} + \epsilon_2/4$ along with the condition $|p_{s+t,0} - p_{s+t,1}| \ge \epsilon_2$ together imply $|\hat{p}_{s+t,0} - \hat{p}_{s+t,1}| \ge \epsilon_2/2$, we obtain

$$\Pr[|\hat{p}_{s+t,0} - \hat{p}_{s+t,1}| \ge \epsilon_2/2] \ge 1 - 4\exp(-N_2\epsilon_2^2/8). \tag{12}$$

For N_2 about ϵ_2^{-2} times a constant, the difference $|\hat{p}_{s+t,0} - \hat{p}_{s+t,1}|$ will be at least $\epsilon_2/2$ with probability almost 1. Keeping this in mind, henceforth we will assume (see Remark 8 below)

$$N_2 = \tilde{O}(\epsilon_2^{-2}). \tag{13}$$

For simplicity, we also assume that for N_2 given by (13), the following holds.

$$\text{If } s+t \text{ is good, then } |\hat{p}_{s+t,0} - \hat{p}_{s+t,1}| \ge \epsilon_2/2 \text{ holds with probability 1.} \tag{14}$$

Given a good $s+t$, we say that $\mathbf{z}$ is good if $|\hat{p}_{s+t,\mathbf{z},0} - \hat{p}_{s+t,\mathbf{z},1}| \ge \epsilon_2/4$ holds. Using Proposition 10 in Appendix B, the probability of a good $\mathbf{z}$ (for a good $s+t$) is at least $\epsilon_2/4$. Changing the error distribution from $D_{\mathfrak{r}}^{\ell N_2}$ to $D_{r'}^{\ell N_2}$, the probability of a good $\mathbf{z}$ under $D_{r'}^{\ell N_2}$ (i.e., one for which $|\hat{\mathfrak{p}}_{s+t,\mathbf{z},0} - \hat{\mathfrak{p}}_{s+t,\mathbf{z},1}| \ge \epsilon_2/4$ holds) is at least $\epsilon_2^2/(256nN_2\ell)^{1/2}$ (see Proposition 14 in Appendix B). So the probability of a good pair $(s+t, \mathbf{z})$ where $\mathbf{z}$ follows $D_{r'}^{N_2\ell}$ is at least $\epsilon_1\epsilon_2^2/(256nN_2\ell)^{1/2}$. Consequently, if N_1 is around $(256nN_2\ell)^{1/2}/(\epsilon_1\epsilon_2^2)$, then with probability exponentially close to 1 a good tuple will be encountered in one of the iterations of the outer loop.

Type-2 failure can occur in two ways. The first way is that in none of the N_1 iterations, a good tuple is obtained. The second way is that for a good tuple, the condition $|\hat{\mathfrak{p}}_{s+t,\mathbf{z},0} - \hat{\mathfrak{p}}_{s+t,\mathbf{z},1}| \ge \epsilon_2/4$ does not hold. The above analysis shows that the probability of either of these errors is exponentially small.

The number of times $\mathcal{A}_7$ calls $\mathcal{D}_1$ is N_1N_2 which is about $(256n\ell)^{1/2}N_2^{3/2}/(\epsilon_1\epsilon_2^2)$ and the number of samples of $A_{s,r}^j$ required by $\mathcal{A}_7$ is $N_1N_2\ell$. As mentioned above N_2 is about ϵ_2^{-2}.

Remark 8. *Ignoring Type-2 failures, the probability that $\mathcal{A}_7$ fails is given by the probability of Type-1 failure. As shown above, this probability is at most $4N_1\exp(-N_2\epsilon_2^2/32)$. In the complete reduction of K-SIVP$_\gamma$ to* ring-VWDLWE$^i_{q,\le\alpha}$*, let N_3 be the number of times $\mathcal{A}_7$ is called. Then the probability that any of these calls fails is at most $4N_1N_3\exp(-N_2\epsilon_2^2/32)$. This shows that the value of N_2 should be about $32L\epsilon_2^{-2}$ where, $L = \ln N_1 + \ln N_3$. Since N_2^2 appears in the number of oracle calls, the factor $32L$ contributes $2^{10}L^2$ to the tightness gap. For practical values of the parameters, this can be significant.*

In view of the above, we have the following result.

Proposition 6. *The number of times $\mathcal{A}_7$ calls $\mathcal{D}_1$ is about $(\epsilon_1\epsilon_2^5)^{-1}\cdot(256n\ell)^{1/2}$. The number of times $\mathcal{A}_7$ calls its LWE oracle is about $(\epsilon_1\epsilon_2^5)^{-1}\cdot(256n)^{1/2}\ell^{3/2}$.*

Remark 9. *We tried various ways to fill in the details of this reduction while maintaining (10) rather than resorting to the much larger $\mathfrak{r}/\alpha$ ratio in (11). The factor (11) came from the need to bound the effect on the probability measure of replacing the distribution $D_{\mathfrak{r}}^{N_2\ell}$ by the actual distribution $D_{r'}^{N_2\ell}$ that governs the error-vectors $\mathbf{z}$ that go into a set of N_2 pairs of lists $(\mathcal{T}', \mathcal{T}'')$ that are input to $\mathcal{D}_2$.*

One promising possibility seemed to be that we could avoid the N_2 term if we were able to consider each inner loop to have an N_2-tuple of lists coming from a fixed set of ℓ samples from $A_{s+t,r}^j$ with fixed t and fixed added error-vectors $\mathbf{f}$, so that the only quantities that vary within an inner loop would be the randomisation h's and the internal coin tosses of the distinguisher.

However, in the case $j = i$ we would need to know that, for any *fixed $s+t$ and for* any *fixed ℓ-tuple of $\mathbf{a}$ and $\mathbf{z}$ coming from the ℓ samples from $A^j_{s+t,r}$, the probabilities of output 1 for $\mathcal{T}'$ and for $\mathcal{T}''$ are equal. This is true if the distributions of h's are the same for the two lists. However, they aren't: for $\mathcal{T}'$ the distribution is $\chi(i-)$ and for $\mathcal{T}''$ it's $\chi(i)$. Speaking less formally, the difference is that the i-th component of $\mathcal{T}'$ is randomised but fixed for all of the N_2 lists, while the i-th component of $\mathcal{T}''$ gets a new randomisation in each new list. One can adopt the heuristic assumption that the distinguisher sees them as indistinguishable, but in that case we don't have a rigorous argument.*

Remark 10. *It is possible to use (10), but in that case we obtain a super-exponential time algorithm. See Remark 24 in Appendix B.*

4.5 Reducing ring-$\mathsf{DLWE}^i_{q,\mathfrak{r}}$ to ring-$\mathsf{DLWE}_{q,\mathfrak{r}}$

The final requirement is a reduction to the decision problem ring-DLWE (see the proof of Lemma 5.14 of [21]). Suppose $\mathcal{D}_2$ is a (δ_1, δ_2)-distinguisher for ring-$\mathsf{DLWE}_{q,\mathfrak{r}}$, i.e., it has advantage at least δ_2 on a fraction δ_1 of the set of possible values of s. We show that there is an $i \in \mathbb{Z}^*_m$, such that $\mathcal{D}_2$ will function as a $(\delta_1/n, \delta_2/n)$-distinguisher $\mathcal{D}_1$ for ring-$\mathsf{DLWE}^i_{q,\mathfrak{r}}$.

We say that s is "useful" if $\mathcal{D}_2$ has advantage at least δ_2 in distinguishing $A_{s,\mathfrak{r}}$ from the uniform distribution over $R_q \times \mathbb{T}$. From the assumption on $\mathcal{D}_2$, the proportion of useful $s \in R^\vee_q$ is at least δ_1. For $s \in R^\vee_q$ and $i \in \mathbb{Z}^*_m \cup \{0\}$, let $p_{s,i}$ be the probability that $\mathcal{D}_2$ outputs 1 on being provided ℓ samples from $A^i_{s,\mathfrak{r}}$. Recall that $A_{s,\mathfrak{r}} = A^0_{s,\mathfrak{r}}$ and that samples from $A^{m-1}_{s,\mathfrak{r}}$ are almost uniformly distributed over $R_q \times \mathbb{T}$. So for a useful s, $|p_{s,0} - p_{s,m-1}| \geq \delta_2$. It then follows that there is at least one $i \in \mathbb{Z}^*_m$ such that $|p_{s,i-} - p_{s,i}| \geq \delta_2/n$. We say that a pair $(s, i) \in R^\vee_q \times \mathbb{Z}^*_m$ is "useful" if the condition $|p_{s,i-} - p_{s,i}| \geq \delta_2/n$ holds. Since the proportion of useful $s \in R^\vee_q$ is at least δ_1, and for each useful s there is at least one i such that (s, i) is useful, it follows that for some i there is a proportion at least δ_1/n of s values such that (s, i) is useful. We then fix such an i, denoted i_0. We define s to be "good" if (s, i_0) is useful. We attempt to use $\mathcal{D}_2$ as our $\mathcal{D}_1$ for the n possible values of i, and when we get to $i = i_0$ we will find that our (δ_1, δ_2)-distinguisher $\mathcal{D}_2$ for ring-$\mathsf{DLWE}_{q,\mathfrak{r}}$ also functions as a $(\delta_1/n, \delta_2/n)$-distinguisher for ring-$\mathsf{DLWE}^i_{q,\mathfrak{r}}$.

Proposition 7. *Suppose $\mathcal{D}_2$ is a (δ_1, δ_2)-distinguisher for ring-*$\mathsf{DLWE}_{q,\mathfrak{r}}$*. Then there is some $i \in \mathbb{Z}^*_m$ such that there is an (ϵ_1, ϵ_2)-distinguisher $\mathcal{D}_1$ for ring-*$\mathsf{DLWE}^i_{q,\mathfrak{r}}$*, where $\epsilon_1 = \delta_1/n$ and $\epsilon_2 = \delta_2/n$.*

Since $\epsilon_2 = \delta_2/n$, the value of N_2 given by (13) becomes

$$N_2 \quad = \quad \tilde{O}(n^2\delta_2^{-2}). \tag{15}$$

Remark 11. *Proposition 7 does not show how to choose an i for which $\mathcal{D}_1$ exists. The argument showing existence of i does not help determine the value of i. We can repeat the full reduction of K-*SIVP_γ *to ring-*$\mathsf{DLWE}_{q,\mathfrak{r}}$ *for each possible i, and among the responses we select the set of independent vectors that has the smallest maximum length. Since there are n possibilities, this introduces a factor of n into the tightness gap.*

4.6 The tightness gap in the search ring-LWE to ring-DLWE reduction

The analysis in this section showed how to construct an algorithm $\mathcal{A}_5$ to solve ring-$\mathsf{LWE}_{q,\leq\alpha}$ using a (δ_1, δ_2)-distinguisher $\mathcal{D}_2$ for ring-$\mathsf{DLWE}_{q,\mathfrak{r}}$. This required going through algorithms $\mathcal{A}_6$ and $\mathcal{A}_7$ and the distinguisher $\mathcal{D}_1$.

First we consider the number of times $\mathcal{A}_5$ calls $\mathcal{D}_2$. By Proposition 4, $\mathcal{A}_5$ calls $\mathcal{A}_6$ a total of n times and by Proposition 5, $\mathcal{A}_6$ calls $\mathcal{A}_7$ a total of q times. By Proposition 6, $\mathcal{A}_7$ calls $\mathcal{D}_1$ about $(\epsilon_1\epsilon_2^5)^{-1} \cdot (256n\ell)^{1/2}$ times. From the discussion in Section 4.5, $\mathcal{D}_1$ is identical to $\mathcal{D}_2$. Thus, the number of times $\mathcal{A}_5$ calls $\mathcal{D}_2$ is about $qn(\epsilon_1\epsilon_2^5)^{-1} \cdot (256n\ell)^{1/2}$.

Next we consider the number of LWE queries made by $\mathcal{A}_5$. By Propositions 4 and 5, the number of LWE queries made by $\mathcal{A}_5$, $\mathcal{A}_6$ and $\mathcal{A}_7$ are equal. By Proposition 6, the number of LWE queries made by $\mathcal{A}_7$ is about $(\epsilon_1\epsilon_2^5)^{-1} \cdot ((256n)^{1/2}\ell^{3/2})$. Hence, the number of LWE queries made by $\mathcal{A}_5$ is also $(\epsilon_1\epsilon_2^5)^{-1} \cdot ((256n)^{1/2}\ell^{3/2})$.

By Proposition 7, we have $\epsilon_1 = \delta_1/n$ and $\epsilon_2 = \delta_2/n$. Using these values of ϵ_1 and ϵ_2 in the above expressions, we obtain the following result.

Theorem 8. *The number of times $\mathcal{A}_5$ calls $\mathcal{D}_2$ is about $qn(\delta_1\delta_2^5)^{-1}n^6 \cdot (256n\ell)^{1/2} \approx qn^{15/2}\ell^{1/2}(\delta_1\delta_2^5)^{-1}$ and the number of LWE queries made by $\mathcal{A}_5$ is about $(\delta_1\delta_2^5)^{-1}n^6 \cdot ((256n)^{1/2}\ell^{3/2}) \approx n^{13/2}\ell^{3/2}(\delta_1\delta_2^5)^{-1}$.*

4.7 The parameters γ and q

In order to have confidence in the security of a ring-DLWE-based cryptosystem, we want to be sure that unless (δ_1, δ_2) is negligible, there is no efficient (δ_1, δ_2)-distinguisher for ring-$\mathsf{DLWE}_{q,\mathfrak{r}}$, where q (the modulus) and $\mathfrak{r}$ (the distribution width) are parameters of our cryptosystem. In particular, we don't want there to be an efficient distinguisher with advantage $2^{-\beta_2}$ unless β_2 is fairly large. We want the K-SIVP_γ problem that reduces to ring-$\mathsf{DLWE}_{q,\mathfrak{r}}$ with this choice of δ_1, δ_2 to be hard. Below we investigate this question using the value $n = 2^{10}$, which has been given as a parameter for some proposed cryptosystems [1, 6].

According to Theorem 3, the approximation factor γ in the SIVP is $\tilde{O}$ of the following expression:

$$\frac{\sqrt{n}}{\alpha} \quad = \quad \frac{\sqrt{n}}{\mathfrak{r}} \left(\frac{nN_2\ell}{\ln(nN_2\ell)} \right)^{1/4} \tag{16}$$

by (11). Note that the parameter δ_1 does not have any effect on γ. As explained in the last paragraph of §3.1 of [21], in the case of a cyclotomic ideal lattice with n a power of 2 we need $\mathfrak{r}$ to be bounded above by $O(\sqrt{\log n/n})$, or else the distribution will be statistically indistinguishable from uniform and no distinguisher will be possible.

Using (15) and the bound $O(\sqrt{\log n/n})$ on $\mathfrak{r}$ and ignoring log factors and constants, from (16) we have

$$\gamma \quad > \quad n(nN_2\ell)^{1/4} > n(n\ell n^2\delta_2^{-2})^{1/4} = n^{7/4}\ell^{1/4}\delta_2^{-1/2} > n^{7/4}\delta_2^{-1/2}. \tag{17}$$

For example, choosing $n = 2^{10}$ and $\delta_2 = 2^{-\beta_2}$, we find that $\gamma > 2^{(35+\beta_2)/2}$.

Now for $\gamma = 2^k$ the fastest classical algorithm known that solves SVP_γ (and hence also solves K-SVP_γ and its equivalent K-SIVP_γ) has running time $2^{\tilde{\theta}(n/k)}$ where $\tilde{\theta}$ suppresses a log factor [26]. We clearly want $2^{n/k}$ to be large. From $\gamma = 2^k > 2^{(35+\beta_2)/2}$, we have $k > (35 + \beta_2)/2$. Suppose we are considering 128-bit security. If we are extra cautious, then we will choose $\beta_2 = 128$; if we are less cautious, then we may choose $\beta_2 = 50$; and if we are not particularly risk-averse we might choose $\beta_2 = 25$. The corresponding lower bounds on γ and upper bounds on $2^{n/k}$ are shown in Table 1. None of these values inspire confidence in the hardness of K-SIVP_γ. In particular, the approximation factors γ are very large, and the running times $2^{n/k}$ are too small.

In addition, a practicality issue arises when we consider the modulus q. A condition for the reductions is that $q\alpha > 2\omega(\sqrt{\ln n})$. Using (11), we obtain

$$q \quad > \quad \frac{2\omega(\sqrt{\ln n})}{\alpha} = \frac{2\omega(\sqrt{\ln n})}{\mathfrak{r}} \left(\frac{nN_2\ell}{\ln(nN_2\ell)} \right)^{1/4}. \tag{18}$$

Again ignoring constants and log-terms and using $\mathfrak{r} < O(\sqrt{\log n/n})$, we have

$$q \quad > \quad n^{5/4}\ell^{1/4}\delta_2^{-1/2} > n^{5/4}\delta_2^{-1/2}. \tag{19}$$

With our values $n = 2^{10}$, $\delta_2 = 2^{-\beta_2}$ we find that $q > 2^{(25+\beta_2)/2}$. The lower bounds for q corresponding to $\beta_2 = 128, 50$ and 25 are shown in Table 1. The cryptosystem would be quite inefficient with these values of the modulus.

δ_2	2^{-128}	2^{-50}	2^{-25}
$\gamma = 2^k$	$2^{81.5}$	$2^{42.5}$	2^{30}
$2^{n/k}$	2^{12}	2^{24}	2^{34}
q	2^{76}	2^{37}	2^{34}

Table 1: For $n = 2^{10}$ the lower bounds on γ and upper bounds on $2^{n/k}$ along with lower bounds on q.

Remark 12. *The NIST-PQC proposals SABER and Kyber are based on module lattices. In §7.2 we consider the reduction for module lattices. The designers of Kyber chose parameters based on ideal lattices of the same dimension as that of module lattices. (See the first paragraph on the second page of [7].) We make the same assumptions for SABER. There are several variants of SABER and Kyber, and the highest security variant for both sets the dimension $n = 1024$. Irrespective of the dimension, the values[12] of q for SABER and Kyber are 2^{13} and 3329 respectively. These values are much lower than the values of q in Table 1.*

5 The tightness gap in the K-SIVP$_\gamma$ to ring-DLWE$_{q,\mathfrak{r}}$ reduction

The tightness gap for the entire reduction is the number of times $\mathcal{A}_0$ calls $\mathcal{D}_2$. This is given by the product of the number of times $\mathcal{A}_0$ calls $\mathcal{A}_5$ and the number of times $\mathcal{A}_5$ calls $\mathcal{D}_2$. The former is given by Theorem 3 to be $(2n + (\log n)/2)n^4 \cdot N$, where N is the number of LWE queries made by $\mathcal{A}_5$, and from Theorem 8 the latter is about $qn^{15/2}\ell^{1/2}(\delta_1\delta_2^5)^{-1}$. Also from Theorem 8, N is about $n^{13/2}\ell^{3/2}(\delta_1\delta_2^5)^{-1}$. Remark 11 shows that we have an additional factor of n in the number of times $\mathcal{A}_0$ calls $\mathcal{D}_2$.

Based on the above analysis, the concrete version of the complete reduction of K-SIVP$_\gamma$ to ring-DLWE$_{q,\mathfrak{r}}$ is given by the following theorem, which corresponds to Theorem 3.6 of [21].

Theorem 9. *Let K be the m-th cyclotomic number field having degree $n = \varphi(m)$ and $R = \mathcal{O}_K$ be its ring of integers. Let $\mathfrak{r}$ be a positive real number bounded from above by $O(\sqrt{\log n/n})$. Let $\delta_1, \delta_2 \in (0, 1]$. Let q be a prime greater than 2 such that $q \equiv 1 \bmod m$ and $q > (2\omega(\sqrt{\ln n})/\mathfrak{r}) \cdot (nN_2\ell/\ln(nN_2\ell))^{1/4}$, where N_2 is defined in the course of the proof and has magnitude $\tilde{O}(n^2/\delta_2^2)$, and ℓ is a positive integer. Suppose there is a (δ_1, δ_2)-distinguisher $\mathcal{D}_2$ which solves ring-DLWE$_{q,\mathfrak{r}}$ given ℓ samples. Then there is a quantum algorithm $\mathcal{A}_0$ requiring approximately $3n^2$ logical qubits to solve K-SIVP$_\gamma$, where $\gamma = \tilde{O}(n^{5/4}\ell^{1/4}/(\mathfrak{r}\delta_2^{1/2}))$. The number of times $\mathcal{A}_0$ calls $\mathcal{D}_2$ is about*

$$(2n + (\log n)/2)n^4 \cdot n^{13/2}\ell^{3/2}(\delta_1\delta_2^5)^{-1} \cdot qn^{15/2}\ell^{1/2}(\delta_1\delta_2^5)^{-1} \cdot n \approx qn^{20}\ell^2 \cdot (\delta_1\delta_2^5)^{-2}. \tag{20}$$

Remark 13. *In arriving at the tightness gap given by (20), we have tried to lower the tightness gap as best we could. Without the particular improvement in our analysis over that of [21] that is noted in Remark 23 of Appendix B, the tightness gap would be $qn^{35}\ell^7 \cdot (\delta_1\delta_2^{10})^{-2}$.*

From the point of view of practical cryptography, it is of interest to consider the tightness gap G for practical values of n. Let us consider $n = 2^{10}$ as in §4.7. Suppose $\delta_1 = 2^{-\beta_1}$, $\delta_2 = 2^{-\beta_2}$. Then q is at least $2^{(25+\beta_2)/2}$ and for simplicity we take $q = 2^{(25+\beta_2)/2}$. Further, following the suggestion in the second paragraph on page 7 of [21], we take $\ell = O(1)$. Then the tightness gap $G = 2^{(425+4\beta_1+21\beta_2)/2}$. Suppose we take $\beta_1 = 0$ and as in §4.7 we consider three values of β_2, namely 128, 50 and 25. The corresponding values of the gap G are $2^{1556.5}$, $2^{737.5}$ and 2^{475} respectively. One may repeat the calculation using other values of β_1 and β_2; for example, taking $\beta_1 = \beta_2 = 128$, the value of G is $2^{1812.5}$.

[12] Downloaded from https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions on February 28, 2022.

To interpret Theorem 9 let's recall how one determines whether a security reduction from a problem $\mathcal{Q}$ (a problem that's believed to be hard) to a problem $\mathcal{P}$ (the problem that our cryptosystem is based on) with tightness gap G provides an assurance of security for given parameters. For us $\mathcal{Q}$=K-SIVP_γ and $\mathcal{P}$=ring-$\mathsf{DLWE}_{q,\mathfrak{r}}$. Suppose that the fastest known algorithm to solve $\mathcal{Q}$ with our parameters has running time T_2. Suppose $\mathcal{P}$ can be attacked by an algorithm taking time T_1. Using the security reduction, we have a second algorithm for $\mathcal{Q}$ that takes time GT_1. We now make the reasonable assumption that this second algorithm will not set a new record for speed in solving $\mathcal{Q}$, and hence $GT_1 \geq T_2$, and so $T_1 \geq T_2/G$. If T_2 is exponential in the parameters and G is fairly small, we can feasibly choose the parameters so that $T_2/G \geq 2^{128}$. In this way practice-oriented provable security can provide convincing evidence of security against mathematical attacks, i.e., attacks that solve $\mathcal{P}$.

If we carry this out for Theorem 9 with the practical value $n = 2^{10}$, we find that the security reduction is worthless as an assurance of security. Since K-SIVP_γ has not yet been extensively studied, let's take the value of T_2 from a harder problem that has been investigated at length, namely SVP. Of course, K-SVP_γ, which is equivalent to K-SIVP_γ, is presumably easier than exact-SVP for general lattices, and so we're likely to overestimate T_2. According to [19], the fastest classical algorithms have heuristic running time $2^{0.337n+o(n)}$ and the fastest quantum algorithms have heuristic running time $2^{0.286n+o(n)}$. Taking $n = 1024$ as before and assuming that the $o(n)$ term doesn't add more than 50 to the exponent, we'll take $T_2 = 2^{395}$ for classical SVP-algorithms and $T_2 = 2^{343}$ for quantum SVP-algorithms. We won't worry about the fact that the T_2 for K-SVP_γ is probably much less, and we'll use the smallest value obtained above for the gap G. With the classical and quantum values we're using for T_2, we get the following lower bounds for the time T_1 needed to break ring-$\mathrm{DLWE}_{q,\mathfrak{r}}$:

$$T_1 > 2^{-80} \ \text{(classical)}; \qquad T_1 > 2^{-132} \ \text{(quantum)}. \tag{21}$$

So the theorem gives us no assurance at all.

Remark 14. *The negative exponents in the lower bounds (21) would be even more extreme if we chose a more realistic estimate for T_2, such as $2^{n/k}$ (see Table 1) and a more cautious value for β_2, say $\beta_2 = 128$. On the other hand, we could get a reasonable lower bound for T_1 for* ring-$\mathsf{DLWE}_{q,\mathfrak{r}}$ *by increasing n, just as F. Gates [14] did for Regev's reductions for general lattices. Because the tightness gap is so large, n would need to be significantly larger than it was for Gates. For example, if $n \approx 2^{17.5}$, then we have*

$$q > 2^{85}, \qquad \gamma = 2^k > 2^{94}, \qquad n/k < 1970, \qquad G \approx 2^{1715},$$

leading to a lower bound for T_1 of approximately 2^{255}.

However, there are two difficulties with choosing n so large. In the first place, the efficiency advantage of ideal lattices would be lost if one has to use lattices of dimension $> 185,000$. In the second place, the quantum part of the reduction, which requires at least $3n^2$ logical qubits, becomes even farther removed from what can reasonably be expected to be feasible. The quantum part would need 10^{11} logical qubits, roughly 20 million times as many as Shor's algorithm needs to factor a 2048-bit RSA modulus. Using a rough comparison with Shor's algorithm (as in (24) below), we estimate that the number of physical gates required would be about 2^{116}. There's a steep price to be paid for significantly increasing n.

5.1 Comparison to the tightness gap in the reduction from SIVP to DLWE for general lattices

The tightness gap of the reduction from SIVP to DLWE for general lattices in [31] was analysed in [10, 35, 14]. All of these prior works overlooked certain aspects of the tightness gap. Taking these into consideration increases the previous estimates. On the other hand, it is possible to reduce the gap by adjusting certain parameter choices in the reduction in [31], which was not done in [10, 35, 14]. We first obtain a more accurate estimate of the tightness gap in [31] and then compare this estimate to the tightness gap in [21]. In the description below, n is

the dimension of the underlying lattice and q is the modulus of the LWE problem.

Tightness gap in the reduction of SIVP to search-LWE. An algorithm $\mathcal{B}_0$ to solve SIVP can be constructed using an algorithm $\mathcal{B}_1$ to solve DGS with a tightness gap of n^3 as in §3.1. An algorithm $\mathcal{B}_1$ to solve DGS can be constructed using an algorithm $\mathcal{B}_5$ to solve search-LWE using algorithms $\mathcal{B}_2$, $\mathcal{B}_3$ and $\mathcal{B}_4$ as intermediate algorithms. The description of $\mathcal{B}_1$ is similar to the description of $\mathcal{A}_1$ given in §3.2. $\mathcal{B}_1$ first prepares a list of I DGS samples of width large enough so that it can do so without invoking the LWE oracle. Then it goes through a loop over $i_0 = 2n + \lceil (\log n)/2 \rceil$ iterations. In each iteration, it updates the list of I samples with a list of another I samples of width at most half of the previous width. This is done by calling a quantum algorithm $\mathcal{B}_2$ which in turn requires the reverse of a BDD solver $\mathcal{B}_3$. The BDD solver $\mathcal{B}_3$ is constructed using a restricted kind of BDD solver $\mathcal{B}_4$. This special BDD solver $\mathcal{B}_4$ uses the LWE solver $\mathcal{B}_5$.

$\mathcal{B}_1$ calls $\mathcal{B}_2$ a total of $i_0 I \approx nI$ times; $\mathcal{B}_2$ calls the reverse of $\mathcal{B}_3$ once; $\mathcal{B}_3$ calls $\mathcal{B}_4$ a total of n times; and the number of times $\mathcal{B}_4$ calls $\mathcal{B}_5$ is a constant multiple of I^2. Here I is the number of LWE samples required by $\mathcal{B}_5$. So the total tightness gap in the reduction of SIVP to search-LWE is $n^5 I^3$.

Tightness gap in the reduction of search-LWE to average-case DLWE. Algorithm $\mathcal{B}_5$ to solve search-LWE can be constructed using an algorithm $\mathcal{B}_6$ to solve worst-case DLWE with a tightness gap of qn. Both $\mathcal{B}_5$ and $\mathcal{B}_6$ require the same number of LWE samples. Algorithm $\mathcal{B}_6$ can be constructed using a (δ_1, δ_2)-distinguisher $\mathcal{D}$ to solve average-case DLWE. The tightness gap of this reduction is $I_1 I_2$ and the number of LWE samples required by $\mathcal{B}_6$ is $I_1 I_2 \ell$, where I_1 and I_2 are constant multiples of δ_1^{-1} and δ_2^{-2} respectively, and ℓ is the number of LWE samples required by $\mathcal{D}$. So the number I of LWE samples required by $\mathcal{B}_5$ is $I_1 I_2 \ell$.

Overall tightness gap. Combining the above two tightness gaps, the overall tightness gap comes out to be

$$qn^6 \ell^3 \cdot (\delta_1 \delta_2^2)^{-4}. \tag{22}$$

Following the discussion in Section 5 of [31], q is a prime between n^2 and $2n^2$ and $\ell = \tilde{O}(n)$. Taking q to be about n^2 and ℓ to be about n, the expression given by (22) reduces to

$$n^{11} \cdot (\delta_1 \delta_2^2)^{-4}. \tag{23}$$

Remark 15. *The above estimate of the tightness gap improves upon the analysis in [31] in the following two ways.*

1. *In [31], the number of times $\mathcal{B}_4$ calls $\mathcal{B}_5$ was taken to be nI^2 so that the failure probability is at most 2^{-n}. For a concrete analysis, it is sufficient to consider the number of times $\mathcal{B}_4$ calls $\mathcal{B}_5$ to be a constant multiple of I^2. Certainly with $n = 1024$ insisting on a 2^{-1024} failure rate would be overkill.*

2. *In [31], I_1 and I_2 were taken to be n/δ_1 and n/δ_2^2 respectively. This ensures that the failure probabilities of the worst-case to average-case reduction are asymptotically zero. The choices, however, are also an overkill. For purposes of concrete analysis it is sufficient to take I_1 and I_2 to be constant multiples of δ_1^{-1} and δ_2^{-2} respectively with constants that, for practical values of n such as 1024, are much less than n.*

The effect of the above two points is that the tightness gap given by (22) is lower by a factor of n^3 compared to the value that would be obtained by following the analysis in [31].

Remark 16. *The two previous works [10, 14] estimated that $\mathcal{B}_4$ calls $\mathcal{B}_5$ n times and overlooked the factor I^2, while this factor was considered in [35]. The fact that for a concrete analysis it is sufficient to consider I^2 and not nI^2 was overlooked in [35]. All the three works [10, 35, 14] considered I to be a polynomial in n and overlooked the fact that $I = I_1 I_2 \ell$. Further, following [31] all these three works considered I_1 and I_2 to be n/δ_1 and n/δ_2^2 respectively.*

The tightness gap in the reduction from SIVP to search-LWE for the reductions in [31] and [21] are $n^5 I^3$ and $n^5 N$ respectively, where N is the number of samples required by the ring-LWE solver. Treating I and N as having similar values, the tightness gap of the reduction for general lattices is greater by a factor of about I^2. This is due to the fact that the definition of the LWE problem in [31] requires the error to follow a fixed width Gaussian distribution. Since the width of the DGS samples is not known, $\mathcal{B}_4$ has to incrementally add errors so that for some increment the error distribution is negligibly far from the error distribution expected by the LWE solver. This step results in the factor I^2 arising in the tightness gap.

We compare the tightness gap of the reduction from approximate SIVP to DLWE for ideal lattices with that for general lattices. Taking $\delta_1 = 2^{-\beta_1}$ and $\delta_2 = 2^{-\beta_2}$, for $n = 1024$, the expression in (23) is $2^{110+4\beta_1+8\beta_2}$. For $\beta_1 = 0$ and $\beta_2 = 128, 50$ and 25, the values of the tightness gap are 2^{1134}, 2^{510} and 2^{310} respectively. For $\beta_1 = \beta_2 = 128$, the tightness gap[13] is 2^{1646}. The tightness gap we obtain for ideal lattices is much more than the tightness gap for general lattices.

Remark 17. *That the tightness gap for ideal lattices is even larger than for general lattices is particularly troubling because the gap is between the problem of cryptographic interest and a problem that is probably easier for ideal lattices than for general lattices.*

Remark 18. *Increasing the value of n to compensate for the tightness gap in [31] was considered in [14]. The above analysis shows that the estimate of the tightness gap in [14] was off the mark and so the suggested values of n to compensate for the gap are also off the mark. Further, as explained above, increasing the value of n to compensate for the tightness gap is futile, since the number of logical qubits grows quadratically in n, making an already infeasible quantum circuit even worse.*

6 Problems with the quantum part of the reduction

The quantum part of the reduction in [21] is largely taken from [31]. In [31] Regev writes that "This article is almost entirely classical. In fact, quantum is needed only in one step in the proof of the main theorem." This is true from the perspective of readers who are trying to understand the description of the reduction. From a pedagogical standpoint, readers who have difficulty following the construction of a quantum state — in this case by applying unitary operators in a 2^N-dimensional Hilbert space with $N > 3 \times 10^6$ — can console themselves with having understood the vast majority of the steps in the security reduction.

However, Regev's statement could also be interpreted as suggesting that the quantum aspect has only a minor effect on the feasibility of the reduction, provided that one assumes that quantum computers scaled to break RSA and ECC will be possible (which is, after all, a motivation for the development of lattice-based cryptography). That could not be further from the truth. In reality, the vast majority of the reduction steps must occur within a quantum computer, and this raises a multitude of questions about feasibility.

The reduction is divided into a sequence of ten algorithms

$$\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_7, \mathcal{D}_1, \mathcal{D}_2.$$

The quantum part starts with $\mathcal{A}_2$. At a key point (the first paragraph of the proof of Lemma 3.14 in [31]) the quantum algorithm has a state that is a linear combination of roughly 2^{3n^2} terms, each involving two entangled registers. The algorithm needs to "erase" the first of these entangled registers, which means "uncomputing" a closest vector in each summand. This is done by converting an algorithm for the closest vector problem (denoted CVP in [31] and BDD in [21]) into a quantum circuit and then reversing the circuit. Thus, the entire circuit for the rest of the reduction from $\mathcal{A}_2$ to $\mathcal{D}_2$ has to be incorporated (after being reversed) into the quantum computation. As a result, as we'll see, the burden on the quantum computer is many orders of magnitude

[13] This was the case considered in [10, 35], where the tightness gap was estimated to be 2^{524}.

greater than it would be for Shor's algorithm to break RSA_{2048}. Most obviously, all of the operations in the sequence of eight algorithms from $\mathcal{A}_2$ to $\mathcal{A}_7$, $\mathcal{D}_1$, $\mathcal{D}_2$ are quantum operations, and according to [38, pp. 96-97] a quantum operation can be expected to cost at least 2^{10} times as much (and possibly up to 2^{50} times as much) as a classical operation.

6.1 The number of logical qubits

The quantum algorithm $\mathcal{A}_2$ is based on Lemma 3.14 of [31], which shows that $n \log R$ logical qubits are required for an ideal $\mathcal{I}$, where R is an integer which is at least $2^{3n}\lambda_n(\mathcal{I})$. Since $\lambda_n(\mathcal{I})$ is generally polynomial in n, it follows that the number of logical qubits required is about $3n^2$. For $n = 2^{10}$ about 3 million logical qubits will be required. In comparison, factoring a 2048-bit RSA modulus requires roughly 4000 to 5000 logical qubits.

6.2 Circuit size

The size of the logical circuit — that is, neglecting error correction — depends on the length and complexity of the algorithm and on the number of logical qubits of input. The algorithms from [21] surveyed in this paper form a complicated interlocking sequence of computations. In addition, the number of qubits for a 1024-dimensional lattice is about 750 times the number of qubits in Shor's algorithm to break RSA_{2048}. The number of gates in the quantum part will be many times greater than in Shor's algorithm, and the circuit depth will also be much greater.

6.3 The number of physical gates

A basic issue in estimating the resources needed for a quantum computation is the need for error correction in order to cope with quantum decoherence. This means that the number of physical gates must be many times the number of logical gates. According to the definitive reference on quantum computing by M. Nielsen and I. Chuang [24], much progress in quantum error correction has been made over the years, and there are reasons for optimism. The main reason cited by Nielsen and Chuang for confidence that the need for error correction is not an insurmountable obstacle is the "threshold theorem." That theorem states that as long as the error probability at a gate is below a certain threshold level, an arbitrarily complicated quantum circuit can be transformed into a physical circuit with negligible error, where the ratio of the number g_p of physical gates to the number g_ℓ of logical gates is polynomial in $\log(g_\ell)$.

We mentioned that the number of logical gates depends on the complexity of the algorithm and, for a given algorithm, on the number of qubits of input. Shor's algorithm is simple. It consists of two computations: a modular exponentiation (or a point multiple in the case of ECC) and a quantum Fourier transform. Nevertheless, the quantum resources needed to apply it to find elliptic curve discrete logarithms or factor integers of cryptographic interest are considerable. Two papers by researchers at Microsoft [33, 16] give concrete estimates for the number of qubits and the physical circuit size (number of Toffoli gates) needed to break RSA and ECC using Shor's algorithm. Solving the Elliptic Curve Discrete Log Problem on an elliptic curve over an n-bit prime field can be done with $9n + 2\lceil \log n \rceil + 10$ qubits and $448n^3 \log n + 4090n^3$ Toffoli gates; for $n = 256$ this means 2330 qubits and roughly 2^{37} Toffoli gates. Factoring an n-bit RSA modulus can be done with $2n + 2$ qubits and $64n^3 \log n + O(n^3)$ Toffoli gates; for $n = 2048$ this means 4098 qubits and roughly 1.5×2^{42} Toffoli gates. In both cases the number of Toffoli gates is roughly proportional to $n^3 \log(n)$, and since n is essentially proportional to the number of qubits, the circuit size also grows proportionally to a log term times the cube of the number of qubits.

Let's try to very roughly extrapolate from integer factorization to the quantum security reductions in [31] and [21] for a 1024-dimensional lattice. Ignoring the $\log(n)$ term and also the fact that the security reductions are far more complicated than Shor's algorithm, we can derive a very rough lower bound for the number of physical gates needed to carry out the security reduction by multiplying the number of Toffoli gates for 2048-bit integer

factorization by the cube of the ratio of the number of qubits for the security reduction to the number for Shor's algorithm:

$$(3 \cdot 1024^2/4098)^3 \times 1.5 \times 2^{42} \quad \approx \quad 2^{71} \text{ gates.} \tag{24}$$

Alternatively, suppose we ignore the growth of the poly($\log(g_\ell)$) factor in the threshold theorem and make the rough assumption that the number of physical gates is simply proportional to the size of the logical circuit and that the latter is proportional to length of the algorithm times the square of the number of qubits. Further suppose that the quantum part of the security reduction is at least 25 times more complicated than Shor's lemma. In that case we arrive at a lower bound of about $(3 \cdot 1024^2/4098)^2 \times 25 \times 1.5 \times 2^{42} \approx 2^{66}$ gates.

6.4 Internal memory

Besides the exorbitant cost in Toffoli gates, there's an obstacle that's intrinsic to the structure of the security reductions in [31] and [21]. The book by Nielsen and Chuang, after discussing the threshold theorem and describing some methods of error correction, concludes with several qualifying remarks, one of which points to the important role of interaction with a classical computer. They comment that their earlier discussion "completely neglected the cost of the classical computations and communication that are done during state preparation, syndrome measurement, and recovery. The cost of these could potentially be quite high" (p. 494). We next describe some concerns about whether it's possible in principle for a classical computer to interact with the quantum computer during the "uncomputation" of a nearest vector in the quantum reduction.

Storing intermediate output while waiting for an algorithm to be ready to use it is problematic in quantum computing. Because of quantum decoherence, "No Loitering" signs are ubiquitous in the quantum computing world. Suppose that an algorithm $\mathcal{A}$ recursively produces a sequence of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ that in a second part of the algorithm get processed in the reverse order to produce the final output. If $\mathcal{A}$ is a standalone algorithm, it can be divided into two quantum algorithms $\mathcal{A}'$ and $\mathcal{A}''$ that interact multiple times with a classical computer. The first of these runs n times. Each run outputs the next $\mathbf{v}_i$, which is read and stored by a classical computer that then rewires the quantum circuit using $\mathbf{v}_i$ to prepare the circuit for computing $\mathbf{v}_{i+1}$. The algorithm $\mathcal{A}''$ also runs n times, where each intermediate output is again read by the classical computer, which prepares the circuit for the next run. Note that "outputting" a value to a classical computer amounts to an observation of a quantum state, and that observation destroys the state.

This necessity of breaking up an algorithm into a sequence of sub-algorithms whose execution is terminated in order to allow rewiring is very different from anything one has to worry about in classical computation, where short-term storage is a basic available resource. Even Shor's quantum algorithm, which in essence is very simple, requires a rewiring for each number we want to factor.

Although this rewiring is key to making quantum computation work, there seems to be a fundamental obstacle to such interaction with a classical computer during the course of the quantum part of the security reduction in [21], because the entire quantum computation (actually, quantum "uncomputation") is being applied to each summand in the quantum state constructed by $\mathcal{A}_2$. This means that it's unclear how the recursive steps could be carried out in practice.

The first place where this concern arises is in the reduction from $\mathcal{A}_3$ to $\mathcal{A}_4$. Starting with the input y to $\mathcal{A}_3$, that reduction recursively generates a sequence of b_i that give a sequence of vectors that are closer and closer to the lattice. When the nearest lattice vector can finally be determined by Babai's algorithm, the second part of the algorithm uses that lattice vector along with the b_i's in reverse order to compute the closest lattice vector to y. Thus, the reduction has the same structure as the algorithm $\mathcal{A}$ at the beginning of this subsection, with the crucial difference that the algorithm cannot be carried out after the first interaction with a classical computer collapses the state that the algorithm is being applied to.

The internal storage issue again arises when we consider the reduction from $\mathcal{A}_4$ to $\mathcal{A}_5$. In that reduction $\mathcal{A}_4$ runs its subroutine (oracle) that produces vectors from the distribution $D_{I,r}$. Then each such vector is used,

along with the q-BDD$_{\mathcal{I},\xi}$-input vector y, to create a sample from $A_{s,r}$ that will be used by $\mathcal{A}_5$ to find s, and hence y. The algorithm $\mathcal{A}_5$ presumably needs to have a lot of samples — that is, a lot of approximate equations — before it can get to work finding s. Where are the first samples from $A_{s,r}$ kept while the later ones are being generated?

The same problem arises in the reduction of $\mathcal{A}_5$ to $\mathcal{A}_7$ (via $\mathcal{A}_6$), which entails finding each of the n components of s modulo $\mathfrak{q}_i$, which are then combined by the Chinese Remainder Theorem to find s modulo q. Each component is determined by running the ring-VWDLWE$^i_{q,\leq\alpha}$ solver for up to q different possible values of the residue. The Chinese Remainder Theorem has to wait for all the components to be computed. How are the first components that are computed going to be preserved while the later ones are being determined?

Finally, in the reduction of $\mathcal{A}_7$ to $\mathcal{D}_1$, the former calls ℓ times upon its oracle for $A^j_{s,r}$ and its oracle for $D_{\mathfrak{r}}$ in order to produce two lists of samples $\mathcal{T}'$ and $\mathcal{T}''$, which $\mathcal{D}_1$ then has to distinguish between (in the case when $j = i-$). It's reasonable to assume that $\mathcal{D}_1$ needs to have the first samples that were created still available after receiving the rest of each list. Again we do not see a way to store that data until it is needed.

6.5 Contrast between classical and quantum reductions

In a classical reduction from a problem $\mathcal{Q}$ to a problem $\mathcal{P}$, the main issue of feasibility is the tightness gap, in other words, the running time of the reduction algorithm that solves $\mathcal{Q}$ given an oracle for $\mathcal{P}$ that runs in unit time. Once the tightness gap is computed and one has a reasonable running time estimate for the best available algorithm for the supposedly hard problem $\mathcal{Q}$, one can give key length recommendations based on the guarantees that the reduction gives. In this way one realizes "practice-oriented provable security" [3].

In quantum reductions, on the other hand, it is not always possible to obtain a meaningful security guarantee simply by increasing key length. One must also consider potential obstacles to feasibility that do not arise in the classical case. As we've seen, one such obstacle is the size of the reduction algorithm's physical circuit, which depends on the number of qubits, the complexity of the algorithm, the nature of the tasks performed, and the way error correction is implemented. In addition, the problem of internal storage requires interventions to prepare the circuit for the next stage of the computations. In some cases it is unclear that this is theoretically possible, let alone achievable in practice.

The quantum reduction for lattice-based cryptography, first described in [31], was a remarkable achievement, opening up a new direction in provable security research. At the same time it also opened up new challenges in determining whether or not a security reduction gives a meaningful guarantee.

6.6 Some caveats

Our discussion of the quantum part of the SIVP$_\gamma$-to-DLWE reduction is of necessity informal, imprecise, and speculative. None of us are experts in quantum computation. In particular, we're making the assumption that our analysis of a quantum BDD algorithm with a DLWE-oracle also applies to the reversed circuit — that is, we're assuming that if the BDD algorithm is infeasible because of quantum decoherence, then so is the reverse circuit.

Moreover, even the experts cannot accurately predict how successful physicists and engineers will be in the coming decades in overcoming the formidable obstacles to the development of a large-scale quantum computer. But in any case, after considering the issues of qubit and circuit size, error correction, and internal memory, it should be clear that, even if it is possible in principle to carry out the reduction, the level of quantum scaling required to do so is far, far greater than the amount needed to break RSA and ECC.

6.7 Summary

There are two reasons to doubt the feasibility of the quantum part of the security reduction in [21]. Even for $n = 1024$ the circuit size is many thousands times the circuit size for Shor's algorithm to factor a 2048-bit RSA

modulus. Since the number of qubits grows quadratically with n, the circuit size becomes much greater if one chooses n large enough to compensate for the tightness gap in the reduction. In addition, the quantum algorithm requires almost continual interaction with a classical computer, and we see no way to do this without destroying the quantum state that the algorithm is being applied to.

Remark 19. *Most of the discussion in this section also applies to the quantum part of [31], upon which the quantum part of [21] is based. For example, both quantum algorithms require at least $3n^2$ logical qubits. The concrete analyses of Regev's reduction in [10, 35, 14] did not look at feasibility of the quantum part.*

7 Subsequent works

We consider several works which extend the results in [21].

7.1 Provably secure variant of NTRU

The NTRU cryptosystem [18] was proposed in 1998, and presently a deterministic variant of this cryptosystem is under consideration for standardisation by NIST. There is no known reduction from a worst-case lattice problem to the task of breaking NTRU. This lack of "provable security" was considered to be a shortcoming. To resolve this shortcoming, in [36] a variant of NTRU was proposed for which it was shown that there is a reduction from the approximate ideal-SIVP problem to breaking the security of the proposed variant. The proof in [36] used the reduction from approximate ideal-SIVP to ring-DLWE in (an earlier version of) [21] as a black box. Further, based on an extension of this reduction to all cyclotomic fields in [13], another provably secure variant of NTRU over cyclotomic fields generated by m-th roots of unity with m prime was proposed in [39]. In view of our concrete analysis of the approximate ideal-SIVP to ring-DLWE reduction, for practical values of parameters there is no reason to believe that the provably secure variants of NTRU in [36, 39] provide any better security than the original variant of NTRU.

7.2 Module lattices

Let K be a number field of degree n and R be its ring of integers. For a positive integer d, K^d is a module over R. The embedding σ of K into H extends to an embedding of K^d into H^d, and one may consider lattices in H^d. The dimension of this lattice is nd. As pointed out in [20], module-LWE serves as a "bridge" between ring-LWE and LWE for general lattices. If $d = 1$ we have ring-LWE for an n-dimensional ideal lattice, and if $n = 1$ we have LWE for a d-dimensional general lattice. For reasons of efficiency, in practice d is a small constant; in Kyber [7] and SABER [12] n is taken to be 256 and d is either 2, 3, or 4. Thus, in practice module-LWE is much closer to ring-LWE than to general lattice LWE. Below we will see that the tightness gap for the reduction from approximate module-SIVP to module-DLWE is a little less than for the K-SIVP_γ to ring-DLWE reduction, but considerably more than for the general SIVP_γ to DLWE reduction. This is not surprising.

In [20] it is suggested that module-DLWE may be harder than ring-DLWE. This is a reasonable conjecture. Indeed, if ring-DLWE is much easier than general DLWE, then the difference in hardness between module-DLWE with $d = 2$, 3, or 4 and general DLWE will probably turn out to be a little less.

The module-LWE distribution is the following. Let $\mathbb{T} = H/\sigma(R^\vee)$ as before and $q \geq 2$ be an integer. For $\mathbf{s} \in (R_q^\vee)^d$ and $\mathbf{a} \in R_q^d$, a sample from the module-LWE distribution with parameters q, $\mathbf{s}$ and r is $(\mathbf{a}, \sigma(\langle \mathbf{a}, \mathbf{s} \rangle / q) + \mathbf{e} \bmod \sigma(R_q^\vee))$, where $\mathbf{e}$ is chosen from H following the distribution D_r. The search and decision versions of the module-LWE problem are defined in a manner analogously to those of ring-LWE problem.

Let K be the m-th cyclotomic field having degree n, $d \geq 1$ be an integer, and q be a prime greater than 2 such that $q \equiv 1 \bmod m$. The reduction in [21] from approximate ideal-SIVP to ring-DLWE was generalised to module lattices in [20], where it was shown that approximate module-SIVP reduces to module-DLWE. The

sequence of algorithms in the reduction in [20] is exactly the same as in the reduction in [21]. In this section we let $\mathcal{A}_0$ to $\mathcal{A}_7$ and $\mathcal{D}_1$ and $\mathcal{D}_2$ denote the algorithms for module-SIVP to module-DLWE reduction. Here $\mathcal{A}_0$ solves module-SIVP_γ and $\mathcal{D}_2$ is a (δ_1, δ_2)-distinguisher for the module-DLWE problem with parameters q and $\mathfrak{r}$. Below we briefly summarise the parameters and the tightness gap of the reduction.

The number of times $\mathcal{A}_0$ calls $\mathcal{A}_5$ is about $(nd)^5 \cdot N$, where N is the number of module-LWE samples required by $\mathcal{A}_5$. The number of times $\mathcal{A}_5$ calls $\mathcal{A}_7$ is qnd and the number of module-LWE samples required by $\mathcal{A}_7$ and $\mathcal{A}_5$ are equal. The number of times $\mathcal{A}_7$ calls the (ϵ_1, ϵ_2)-distinguisher $\mathcal{D}_1$ is $N_1 N_2$ and the number of module-LWE samples required by $\mathcal{A}_7$ is $N_1 N_2 \ell$, where $N_1 = (nN_2\ell)^{1/2}/(\epsilon_1\epsilon_2^2)$, $N_2 = \tilde{O}(\epsilon_2^{-2})$ and ℓ is the number of LWE-samples required by $\mathcal{D}_1$. So $N = N_1 N_2 \ell$. The (ϵ_1, ϵ_2)-distinguisher $\mathcal{D}_1$ can be used as a (δ_1, δ_2) distinguisher $\mathcal{D}_2$, where $\epsilon_1 = \delta_1/n$ and $\epsilon_2 = \delta_2/n$. The entire reduction from module-SIVP to module-DLWE needs to be invoked an additional n times due to the reason noted in Remark 11. The overall tightness gap is about

$$qn^{20}d^6\ell^2(\delta_1\delta_2^5)^{-2}. \tag{25}$$

Also, the quantum part of the reduction requires about $(nd)^2$ logical qubits.

For fixed dimension nd the tightness gap in (25) for the module case is less than that in (20) for the ring case by a factor of d^{14}. Thus, for the values $n = 256$ and $d = 4$ – the values given for a high security level for Kyber and SABER – the lower bounds for T_1 in (21) should be replaced by

$$T_1 > 2^{-52} \text{ (classical)}; \qquad T_1 > 2^{-104} \text{ (quantum)}.$$

Remark 20. *The discussion in §6 regarding the quantum part of the reduction of approximate ideal-SIVP to ring-DLWE in [21] also applies to the reduction of approximate module-SIVP to module-DLWE in [20].*

7.3 Arbitrary number fields

A follow-up work [28] (whose latest version is [29]) improved upon the reduction in [21]. The reduction of K-SIVP_γ to ring-$\mathsf{DLWE}_{q,r}$ in [21] holds only for cyclotomic number fields. In [28, 29] this reduction was extended to any number field. Unlike [21], the reduction in [28, 29] does not go through the search ring-LWE problem. Below we discuss the tightness gap in the reduction in [28, 29]. Specifically, we refer to the version in [29].

As in [21], the reduction of K-SIVP_γ to ring-$\mathsf{DLWE}_{q,r}$ in [29] also goes through several steps. The first step is to reduce K-SIVP_γ to K-DGS_Γ and is exactly the algorithm $\mathcal{A}_0$ described earlier that has a tightness gap of n^3. The reduction of K-DGS_Γ to ring-$\mathsf{DLWE}_{q,r}$ is given by a sequence of algorithms. The first of these is the algorithm $\mathcal{A}_1$ described earlier. Briefly, $\mathcal{A}_1$ prepares an initial list of N DGS samples and goes through $i_0 = (2n + (\log n)/2)$ iterations, where in each iteration $\mathcal{A}_1$ invokes the quantum circuit $\mathcal{A}_2$ N times, and in each invocation $\mathcal{A}_1$ provides $\mathcal{A}_2$ with a list of DGS samples and receives in return DGS samples with width reduced by a factor of at least 2. Finally, $\mathcal{A}_1$ returns a sample from the last list that it prepares. The number N of DGS samples is equal to the number of LWE samples required by a distinguisher for the ring-$\mathsf{DLWE}_{q,r}$ problem, and this number in [29] is very different from that of [21]. As before, $\mathcal{A}_1$ calls $\mathcal{A}_2$ a total of $(2n + (\log n)/2)N$ times.

From this point onwards, the reductions in [21] and [29] begin to differ. In [29], $\mathcal{A}_2$ applies the reverse of an algorithm $\mathcal{B}_3$ that solves the Gaussian decoding problem[14]. The construction of $\mathcal{B}_3$ is based on a distinguisher $\mathcal{E}_1$ for the ring-$\mathsf{DLWE}_{q,\leq\alpha}$ problem[15]. The construction of $\mathcal{E}_1$ is based on a (δ_1, δ_2)-distinguisher $\mathcal{D}_2$ for the ring-$\mathsf{DLWE}_{q,\mathfrak{r}}$ problem where the relation between $\mathfrak{r}$ and α is given by (11). The construction of $\mathcal{E}_1$ from $\mathcal{D}_2$ is very similar to the construction of $\mathcal{A}_7$ from $\mathcal{D}_1$ described in §4.4, and the concreteness aspects are exactly the same. In particular, $\mathcal{E}_1$ calls $\mathcal{D}_2$ about $(\delta_1\delta_2^5)^{-1}(n\ell)^{1/2}$ times and the number M_1 of LWE samples required by $\mathcal{E}_1$ is about $(\delta_1\delta_2^5)^{-1} \cdot (n^{1/2}\ell^{3/2})$, where ℓ is the number of LWE samples required by $\mathcal{D}_2$.

[14] For a lattice $\Lambda \subset H$ and a parameter ξ, an instance of the problem is a coset $\mathbf{e} + \Lambda$, where $\mathbf{e}$ is drawn from D_ξ, and the task is to find $\mathbf{e}$. Note that this is essentially our definition of the BDD problem in Section 3.

[15] Formally, it is necessary to consider elliptical Gaussian distribution, but this does not matter for the tightness analysis.

The main technical contribution of [29] is the construction of $\mathcal{B}_3$ from $\mathcal{E}_1$. We will not get into the details of this very complicated construction, and instead we simply identify the tightness gap of this reduction. Let M_2 be the number of times $\mathcal{B}_3$ calls $\mathcal{E}_1$. From the description in [29], we obtain an estimate of M_2. Note that each call to $\mathcal{E}_1$ requires M_1 samples so that the total number of samples required in all the M_2 calls is equal to M_1M_2. This is the value of N, since M_1M_2 DGS samples need to be provided to $\mathcal{B}_3$ so that it can generate the required number of LWE samples.

Let[16] $\kappa = \text{poly}(n) \geq 100n^2M_1$ and $\mu = \text{poly}(\kappa)$. Recall that $n = s_1 + 2s_2$, where s_1 and $2s_2$ are the numbers of real and complex roots respectively of the defining polynomial of the number field. Using $\mathcal{E}_1$, $\mathcal{B}_3$ creates s_1+s_2 oracles, where each of these oracles calls $\mathcal{E}_1$ once (Lemma 6.6 in [29]). Corresponding to these $s_1 + s_2$ oracles, $s_1 + s_2$ algorithms are created, where (ignoring logarithmic factors) each algorithm calls its corresponding oracle about $5 \times 10^{15} \cdot \kappa^6\mu^3$ times, where κ and μ are at least $100n^2M_1$ (first part of the proof of Proposition 4.4 and the proof of Lemma 6.6 of [29]). Each of these $s_1 + s_2$ algorithms is itself called about $2000\kappa^3$ times (second part of the proof of Proposition 4.4 of [29]) For simplicity, we take $\mu = \kappa = 100n^2M_1$ and take $2(s_1 + s_2)$ to be n to obtain the value of M_2 to be about $10^{43} \cdot n^{25}M_1^{12}$.

The overall tightness gap is given by the number of times $\mathcal{A}_0$ calls $\mathcal{D}_2$. This number is about

$$\begin{aligned} n^3 \cdot (2n + (\log n)/2)N \cdot M_2 \cdot (\delta_1\delta_2^5)^{-1}(n\ell)^{1/2} &\approx n^4 \cdot M_1M_2^2 \cdot (\delta_1\delta_2^5)^{-1}(n\ell)^{1/2} \\ &\approx 10^{86} \cdot n^{54} \cdot M_1^{25} \cdot (\delta_1\delta_2^5)^{-1}(n\ell)^{1/2} \\ &\approx 10^{86} \cdot n^{54} \cdot ((\delta_1\delta_2^5)^{-1} \cdot (n^{1/2}\ell^{3/2}))^{25} \cdot (\delta_1\delta_2^5)^{-1}(n\ell)^{1/2} \\ &= 10^{86} \cdot n^{67} \cdot \ell^{38} \cdot (\delta_1\delta_2^5)^{-26}. \end{aligned} \tag{26}$$

Remark 21. *The discussion in §6 regarding the quantum circuit required for the reduction in [21] applies equally to the reduction in [29]*

The estimate of the tightness gap given by (26) shows that from a practical point of view the reduction is completely meaningless. This estimate may also be compared with the estimate of the tightness gap of the reduction in [21] given by (20) and the tightness gap of the reduction in [31] given by (22). While the tightness gap of the reduction in [31] is itself huge, it is lower than the tightness gap of the reduction in [21], and the tightness gap of the reduction in [29] is much much larger than the tightness gaps of the reductions in both [31] and [21].

Remark 22. *It was remarked in [29] (sentence before Definition 4.1) that the authors did not try to optimise the parameters. So it is possible that the estimate given by (26) can be lowered with optimised parameters.*

8 Conclusion

Our main result is that the security reduction in [21] gives no meaningful guarantee of real-world security for three reasons – a huge tightness gap, obstacles to realisation of the quantum part, and an approximation factor for the K-SIVP_γ problem that makes it unlikely that the problem is hard, particularly since the lattice problem has been restricted to a subset of lattices with special geometric and algebraic properties. In addition, the reduction requires a value for the modulus q that is much larger than in proposed implementations.

But we need to qualify this by clarifying what we are not claiming. First, we do not have an actual attack on LWE security, and are not saying that one necessarily exists. Nor are we saying that the tightness gap in the reduction in [21] is intrinsic to the approximate ideal-SIVP_γ and DLWE problems; although we have made every effort to lower the tightness gap, it is certainly possible that a reduction could be constructed with an even lower tightness gap. Similarly, we invite the reader to look for a reduction of ring-LWE to ring-DLWE with $\mathfrak{r}/\alpha$ equal

[16] The constant 100 is from [29].

to $(n\ell/\ln(n\ell))^{1/4}$ as claimed in [21] rather than $(nN_2\ell/\ln(nN_2\ell))^{1/4}$, which was the best we could rigorously justify.

In theory, it would be possible to compensate for the tightness gap by increasing n. However, this is not so simple as for classical reductions because, as remarked in Section 5, it entails a sharp increase in the number of logical qubits and physical gates in the quantum part of the reduction. Although the number of logical qubits is "only" quadratic in n, in the quantum context that rate of growth gives rise to grave doubts about feasibility.

We have no intention of questioning the quality of the work in [21]. The construction of the security reduction involving ten nested sub-algorithms was a true *tour de force*. From a theoretical standpoint, it was a major accomplishment to construct a polynomial time reduction from approximate ideal-SIVP$_\gamma$ to the decision problem that proposed cryptosystems are based on. Unfortunately, the reduction loses its value when viewed from the vantage point of practice-oriented provable security.

The four major problems in the reduction in [21] (the large tightness gap, the large value of the approximation factor, the unrealistic quantum part, and the likelihood that approximate ideal-SIVP is substantially easier than approximate SIVP for general lattices) mean that one cannot have any confidence that the reduction rules out practical mathematical attacks on ring-DLWE based cryptosystems. It seems that similar doubts apply to module-DLWE based systems, although it is reasonable to conjecture that such a system with $n = 256$ and $d = 4$ is a little less vulnerable than a ring-DLWE based system with $n = 1024$.

In 2015 Peikert[17] discussed asymptotic analyses of the security of lattice-based systems and concluded that they ensure the superiority of such systems from the standpoint of security:

> ...worst-case reductions give a hard-and-fast guarantee that the cryptosystem is at least as hard to break as the hardest instances of some underlying problem. This gives a true lower bound on security, and prevents the kind of unexpected weaknesses that have so often been exposed in schemes that lack such reductions.

But the type of security reductions analysed in this paper give no guarantees of real-world security, let alone "hard-and-fast" ones. There is no meaningful "true lower bound" on security. Among the proposals for post-quantum cryptography, the ones that have badly deficient "proofs of security" should not be privileged over those that have arguments for security that are grounded in heuristics and practical analysis.

Acknowledgements

We wish to thank Dan Bernstein and Alfred Menezes for helpful comments on an earlier draft and Ann Hibner Koblitz for editorial corrections and comments. Of course, any opinions or errors in the paper are the sole responsibility of the authors.

References

[1] Erdem Alkim, Léo Ducas, Thomas Pöppelmann, and Peter Schwabe. Post-quantum key exchange - a new hope. Cryptology ePrint Archive, Report 2015/1092, 2015. https://ia.cr/2015/1092.

[2] W. Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296:625–636, 1993.

[3] Mihir Bellare. Practice-oriented provable-security. In *International Workshop on Information Security*, volume 1396 of *Lecture Notes in Computer Science*, pages 221–231. Springer, 1997.

[17] C. Peikert, 19 February 2015 blog posting, http://web.eecs.umich.edu/~cpeikert/soliloquy.html; accessed on March 5, 2022.

[4] Daniel J. Bernstein. Comparing proofs of security for lattice-based encryption. Cryptology ePrint Archive, Report 2019/691, 2019. https://eprint.iacr.org/2019/691.

[5] Daniel J. Bernstein and Tanja Lange. Non-randomness of S-unit lattices. Cryptology ePrint Archive, Report 2021/1428, 2021. https://eprint.iacr.org/2021/1428.

[6] Joppe W. Bos, Craig Costello, Michael Naehrig, and Douglas Stebila. Post-quantum key exchange for the TLS protocol from the ring learning with errors problem. In *IEEE Symposium on Security and Privacy*, pages 553–570. IEEE Computer Society, 2015.

[7] Joppe W. Bos, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Vadim Lyubashevsky, John M. Schanck, Peter Schwabe, Gregor Seiler, and Damien Stehlé. CRYSTALS - kyber: A cca-secure module-lattice-based KEM. In *EuroS&P*, pages 353–367. IEEE, 2018.

[8] Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, pages 575–584. ACM, 2013.

[9] George Casella and Roger L. Berger. *Statistical Inference*. Thomson, 2002.

[10] Sanjit Chatterjee, Neal Koblitz, Alfred Menezes, and Palash Sarkar. Another look at tightness II: practical issues in cryptography. In Raphael C.-W. Phan and Moti Yung, editors, *Paradigms in Cryptology - Mycrypt 2016. Malicious and Exploratory Cryptology - Second International Conference, Mycrypt 2016, Kuala Lumpur, Malaysia, December 1-2, 2016, Revised Selected Papers*, volume 10311 of *Lecture Notes in Computer Science*, pages 21–55. Springer, 2016.

[11] Ronald Cramer, Léo Ducas, and Benjamin Wesolowski. Mildly short vectors in cyclotomic ideal lattices in quantum polynomial time. *J. ACM*, 68(2):8:1–8:26, 2021.

[12] Jan-Pieter D'Anvers, Angshuman Karmakar, Sujoy Sinha Roy, and Frederik Vercauteren. Saber: Module-lwr based key exchange, cpa-secure encryption and cca-secure KEM. In *AFRICACRYPT*, volume 10831 of *Lecture Notes in Computer Science*, pages 282–305. Springer, 2018.

[13] Léo Ducas and Alain Durmus. Ring-LWE in polynomial rings. In *Public Key Cryptography*, volume 7293 of *Lecture Notes in Computer Science*, pages 34–51. Springer, 2012.

[14] Fletcher Gates. Reduction-respecting parameters for lattice-based cryptosystems. https://macsphere.mcmaster.ca/bitstream/11375/24466/2/gates_fletcher_m_finalsubmission2018october_msc.pdf, 2018. Masters Thesis, McMaster University.

[15] Oded Goldreich and Shafi Goldwasser. On the limits of nonapproximability of lattice problems. *J. Comput. Syst. Sci.*, 60(3):540–563, 2000.

[16] Thomas Häner, Martin Roetteler, and Krysta M. Svore. Factoring using $2n + 2$ qubits with Toffoli based modular multiplication. arXiv preprint https://arxiv.org/abs/1611.07995, 2016.

[17] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[18] Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman. NTRU: A ring-based public key cryptosystem. In *Algorithmic Number Theory, Third International Symposium, ANTS-III, Portland, Oregon, USA, June 21-25, 1998, Proceedings*, pages 267–288, 1998.

[19] Thijs Laarhoven, Michele Mosca, and Joop van de Pol. Finding shortest lattice vectors faster using quantum search. *Designs, Codes and Cryptography*, 77(2-3):375–400, 2015.

[20] Adeline Langlois and Damien Stehlé. Worst-case to average-case reductions for module lattices. *Designs, Codes and Cryptography*, 75(3):565–599, 2015.

[21] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. *J. ACM*, 60(6):43:1–43:35, 2013.

[22] D. Micciancio and S. Goldwasser. *Complexity of Lattice Problems: A Cryptographic Perspective*. Springer, 2002.

[23] Daniele Micciancio. Lattice algorithms and applications: Basis reduction. https://cseweb.ucsd.edu/classes/sp14/cse206A-a/lec5.pdf, 2014. Accessed on February 8, 2022.

[24] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2nd edition, 2010.

[25] Chris Peikert. Public-key cryptosystems from the worst-case shortest vector problem: extended abstract. In Michael Mitzenmacher, editor, *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 333–342. ACM, 2009.

[26] Chris Peikert. A decade of lattice cryptography. *Foundations and Trends in Theoretical Computer Science*, 10(4):283–424, 2016. https://doi.org/10.1561/0400000074.

[27] Chris Peikert and Zachary Pepin. Algebraically structured LWE, revisited. In *TCC (1)*, volume 11891 of *Lecture Notes in Computer Science*, pages 1–23. Springer, 2019.

[28] Chris Peikert, Oded Regev, and Noah Stephens-Davidowitz. Pseudorandomness of ring-LWE for any ring and modulus. In *STOC*, pages 461–473. ACM, 2017.

[29] Chris Peikert, Oded Regev, and Noah Stephens-Davidowitz. Pseudorandomness of ring-LWE for any ring and modulus. *IACR Cryptol. ePrint Arch.*, 2017. https://eprint.iacr.org/2017/258, version dated 6 June, 2020.

[30] Oded Regev. Lattices in computer science: Average-case hardness. https://cims.nyu.edu/~regev/teaching/lattices_fall_2004/ln/averagecase.pdf, accessed on February 8, 2022, 2004.

[31] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6):34:1–34:40, 2009.

[32] Oded Regev. The learning with errors problem (invited survey). In *Computational Complexity Conference*, pages 191–204. IEEE Computer Society, 2010.

[33] Martin Roetteler, Michael Naehrig, Krysta M. Svore, and Kristin E. Lauter. Quantum resource estimates for computing elliptic curve discrete logarithms. In *ASIACRYPT (2)*, volume 10625 of *Lecture Notes in Computer Science*, pages 241–270. Springer, 2017.

[34] Palash Sarkar and Subhadip Singha. Classical reduction of gap SVP to LWE: A concrete security analysis. *Advances in Mathematics of Communications*, 2021. https://www.aimsciences.org/article/doi/10.3934/amc.2021004.

[35] Palash Sarkar and Subhadip Singha. Verifying solutions to LWE with implications for concrete security. *Advances in Mathematics of Communications*, 15(2):257–266, 2021.

[36] Damien Stehlé and Ron Steinfeld. Making NTRU as secure as worst-case problems over ideal lattices. In *EUROCRYPT*, volume 6632 of *Lecture Notes in Computer Science*, pages 27–47. Springer, 2011.

[37] Damien Stehlé, Ron Steinfeld, Keisuke Tanaka, and Keita Xagawa. Efficient public key encryption based on ideal lattices. In *ASIACRYPT*, volume 5912 of *Lecture Notes in Computer Science*, pages 617–635. Springer, 2009.

[38] NTRU Prime Risk-Management Team. Risks of lattice KEMs. 2021. https://ntruprime.cr.yp.to/latticerisks-20211031.pdf, accessed on February 18, 2022.

[39] Yang Yu, Guangwu Xu, and Xiaoyun Wang. Provably secure NTRU instances over prime cyclotomic rings. In *Public Key Cryptography (1)*, volume 10174 of *Lecture Notes in Computer Science*, pages 409–434. Springer, 2017.

A Details of the parameters of $\mathcal{A}_2$ and $\mathcal{A}_3$

Algorithm $\mathcal{A}_2$ takes as input an ideal $\mathcal{I}$ and a set of samples from $D_{\mathcal{I},r}$ and returns a sample from $D_{\mathcal{I},r'}$. Algorithm $\mathcal{A}_3$ takes as input a pair $(\mathcal{I}^\vee, y)$, where $\mathcal{I}$ is a fractional ideal of K, $y = x + e$, $x \in \mathcal{I}^\vee$ and $e = \sigma^{-1}(\mathbf{e})$ with $\mathbf{e}$ being chosen according to the distribution D_ξ. Additionally, $\mathcal{A}_3$ also has access to a set of samples from $D_{\mathcal{I},r}$. The parameters[18] r, r' and ξ have to satisfy the following relations.

$$\left.\begin{array}{rcl} r & \geq & \sqrt{2}q \cdot \eta_\varepsilon(\mathcal{I}), \\ r' & = & r \cdot \omega(\sqrt{\ln n})/(\alpha q) > \sqrt{2n}/\lambda_1(\mathcal{I}^\vee), \\ \xi & = & (\alpha q)/(2r \cdot \omega(\sqrt{\ln n})) < \lambda_1(\mathcal{I}^\vee)/(2\sqrt{2n}). \end{array}\right\} \tag{27}$$

Let $\xi' = \xi \cdot \sqrt{2} \cdot \omega(\sqrt{\ln n}) = (\alpha q)/(\sqrt{2}r)$. For $\mathbf{e}$ chosen according to D_ξ, $\|\mathbf{e}\|_\infty \leq \xi'$ except with negligible probability.

The input to Algorithm $\mathcal{A}_1$ is a pair $(\mathcal{I}, r)$ where $r \geq \Gamma(\mathcal{I})$ and $\Gamma(\mathcal{I})$ is given by (6). Recall that $i_0 = 2n + \lceil (\log n)/2 \rceil$ and $r_i = r \cdot (\alpha q/\omega(\sqrt{\ln n}))^i \geq 2^i r$ for $i = 0, \ldots, i_0$ and $\xi_i = (\alpha q)/(2r_i \cdot \omega(\sqrt{\ln n}))$ for $i = 1, \ldots, i_0$. Also, we have $\xi_i' = \alpha q/(\sqrt{2}r_i)$. Note that $r_{i-1} = r_i \cdot \omega(\sqrt{\ln n})/(\alpha q) \leq r_i/2$ for $i = 1, \ldots, i_0$, and $\mathcal{S}_i$ is a list of independent samples from $D_{\mathcal{I},r_i}$, for $i = 0, \ldots, i_0$.

Algorithm $\mathcal{A}_1$ first prepares $\mathcal{S}_{i_0}$. Then for each i going down from i_0 to 1, $\mathcal{A}_1$ gives $\mathcal{A}_2$ the input $\mathcal{I}$ and the set $\mathcal{S}_i$ and receives in return a sample from $D_{\mathcal{I},r_{i-1}}$. This is done N times to prepare the list $\mathcal{S}_{i-1}$. Since the samples $\mathcal{S}_{i_0}, \ldots, \mathcal{S}_1$ are provided as input to $\mathcal{A}_2$, the values $r_{i_0}, \ldots, r_1$ have to satisfy the condition on r given in (27). Further, since $\mathcal{A}_2$ provides samples from $D_{\mathcal{I},r_{i_0-1}}, \ldots, D_{\mathcal{I},0}$ as output, the values $r_{i_0-1}, \ldots, r_0$ have to satisfy the condition on r' given in (27). Algorithm $\mathcal{A}_3$ is invoked on instances $(\mathcal{I}^\vee, y_{i_0}), \ldots, (\mathcal{I}^\vee, y_1)$, where the offsets $e_{i_0}, \ldots, e_1$ in the $y_{i_0}, \ldots, y_1$ are sampled respectively from $D_{\xi_{i_0}}, \ldots, D_{\xi_1}$ so that the values $\xi_{i_0}, \ldots, \xi_1$ have to satisfy the conditions on ξ given in (27). The conditions on r_i and ξ_i are shown below. Before that we show $r_{i_0} \geq 2^{2n}\lambda_n(\mathcal{I})$, so that $\mathcal{A}_1$ can prepare $\mathcal{S}_{i_0}$ directly (i.e., without invoking $\mathcal{A}_2$).

Claim 2.13 of [31] shows that the inequality $\eta_\varepsilon(\mathcal{I}) > 1/\lambda_1(\mathcal{I}^\vee)$ holds for $\varepsilon \leq e^{-\pi}$. Under the assumptions $\varepsilon \leq e^{-\pi}$ and $\alpha < \sqrt{\ln n/n}$, it follows that

$$\Gamma(\mathcal{I}) \;>\; \sqrt{2n}/\lambda_1(\mathcal{I}^\vee). \tag{28}$$

[18] In terms of the notation used in Lemma 4.3, 4.4 and 4.7 of [21], $d' = \xi \cdot \sqrt{2n}$ and $d = \xi'$.

The following computation shows that $r_{i_0} \geq 2^{2n}\lambda_n(\mathcal{I})$.

$$\begin{aligned} r_{i_0} &\geq 2^{i_0} \cdot r \\ &\geq 2^{2n+\frac{1}{2}\log n} \cdot \frac{\sqrt{2}\cdot\omega(\sqrt{\ln n})\cdot\eta_\varepsilon(\mathcal{I})}{\alpha} \quad \text{(using (6))} \\ &\geq \sqrt{n}2^{2n} \cdot \frac{\sqrt{2}\cdot\omega(\sqrt{\ln n})}{\alpha} \cdot \sqrt{\frac{\ln 1/\varepsilon}{\pi}} \cdot \frac{\lambda_n(\mathcal{I})}{n} \quad \text{(using Claim 2.13 of [31][19])} \\ &> 2^{2n}\lambda_n(\mathcal{I}) \quad \text{(using } \omega(\sqrt{\ln n}) > \sqrt{\ln n},\ \alpha < \sqrt{\ln n/n} \text{ and } \varepsilon < e^{-\pi}\text{).} \end{aligned}$$

For $i = 1, \ldots, i_0$, we have the following.

- Since $r \geq \Gamma(\mathcal{I})$, using the definition of $\Gamma(\mathcal{I})$ in (6), we have $r_i \geq r \cdot (\alpha q)/\omega(\sqrt{\ln n}) \geq \sqrt{2}q \cdot \eta_\varepsilon(\mathcal{I})$.
- Since $r_{i-1} \geq r$, and $r \geq \Gamma(\mathcal{I})$, using (28), we have $r_{i-1} > \sqrt{2n}/\lambda_1(\mathcal{I}^\vee)$.
- Noting that $\xi_i = 1/(2r_{i-1})$, using the previous point, we have $\xi_i < \lambda_1(\mathcal{I}^\vee)/(2\sqrt{2n})$.

B Details of the analysis in Section 4.4

We first show the lower bound on the probability of a good $\mathbf{z}$ for a good $s+t$.

Proposition 10. *For a good $s+t$, under the error distribution $D_{\mathfrak{r}}^{\ell N_2}$ the probability that $\mathbf{z}$ is good is at least $\epsilon_2/4$.*

Proof. Since $s+t$ is good, from (14) we have $|\hat{p}_{s+t,0} - \hat{p}_{s+t,1}| \geq \epsilon_2/2$. Without loss of generality, we assume $\hat{p}_{s+t,0} \geq \hat{p}_{s+t,1} + \epsilon_2/2$. Let Z denote the set of all $N_2\ell$-tuples $\mathbf{z}$, and let Y denote the subset of Z consisting of $\mathbf{z}$ such that $\hat{p}_{s+t,\mathbf{z},0} \geq \hat{p}_{s+t,\mathbf{z},1} + \epsilon_2/4$. We claim that Y has measure at least $\epsilon_2/4$. Assume the contrary. We then have

$$\begin{aligned} \hat{p}_{s+t,0} &= \int_Y \hat{p}_{s+t,\mathbf{z},0} D_{\mathfrak{r}}^{\ell N_2}(\mathbf{z}) + \int_{Z\setminus Y} \hat{p}_{s+t,\mathbf{z},0} D_{\mathfrak{r}}^{\ell N_2}(\mathbf{z}) \\ &< \int_Y 1 \cdot D_{\mathfrak{r}}^{\ell N_2}(\mathbf{z}) + \int_Z (\hat{p}_{s+t,\mathbf{z},1} + \epsilon_2/4) D_{\mathfrak{r}}^{\ell N_2}(\mathbf{z}) \\ &\leq \epsilon_2/4 + \hat{p}_{s+t,1} + \epsilon_2/4, \end{aligned}$$

a contradiction. This shows that the probability of $\mathbf{z}$ being good is at least $\epsilon_2/4$. □

Next we consider the effect of changing the error distribution from $D_{\mathfrak{r}}^{\ell N_2}$ to $D_{r'}^{\ell N_2}$. To do this, we introduce a quantity whose logarithm is the Rényi divergence of order 2. Let k be a positive integer. For two probability density functions[20] $P, Q : H^k \to \mathbb{R}_{\geq 0}$, let

$$\mathrm{R}(P||Q) = \int_{H^k} \frac{P(x)^2}{Q(x)} dx. \tag{29}$$

[19] Claim 2.13 of [31] shows that $\eta_\varepsilon(\mathcal{I}) \geq \sqrt{\ln(1/\varepsilon)/\pi} \cdot \lambda_n(\mathcal{I})/n$.

[20] In Claim 5.11 of [21], the density functions are considered to be over $\mathbb{R}^n$. Here we consider density functions over H^k.

function $\mathcal{A}_7$
  for N_1 iterations **do**
    Choose t uniformly from $R_q^\vee$
    $\mathsf{cnt}_0 \leftarrow 0$; $\mathsf{cnt}_1 \leftarrow 0$;
    for N_2 iterations **do**
      Obtain a list $\mathcal{T}$ of ℓ samples from $A_{s,r}^i$
      Choose $\mathbf{f}_1, \ldots, \mathbf{f}_\ell$ independently from $D_{\mathfrak{r}}^\ell$
      Compute $\mathcal{T}'$ and $\mathcal{T}''$ from $\mathcal{T}$, t and $\mathbf{f}_1, \ldots, \mathbf{f}_\ell$
      $\mathsf{cnt}_0 \leftarrow \mathsf{cnt}_0 + \mathcal{D}_1(\mathcal{T}')$; $\mathsf{cnt}_1 \leftarrow \mathsf{cnt}_1 + \mathcal{D}_1(\mathcal{T}'')$
    end for
    $\hat{\mathfrak{p}}_0 \leftarrow \mathsf{cnt}_0/N_2$; $\hat{\mathfrak{p}}_1 \leftarrow \mathsf{cnt}_1/N_2$
    if $|\hat{\mathfrak{p}}_0 - \hat{\mathfrak{p}}_1| \geq \epsilon_2/4$ **then return** $i-$
  end for
  return i
end function.

Table 2: Pseudo-code of algorithm $\mathcal{A}_7$ from distinguisher $\mathcal{D}_1$ described in Section 4.4.

By an abuse of notation, we will also write $\mathrm{R}(D||D')$ to denote $\mathrm{R}(P||Q)$, where D and D' are the distributions corresponding to P and Q respectively. For a measureable subset B of H^k, we have,

$$
\begin{aligned}
(\Pr_D[B])^2 &= \left(\int_B P(x)dx\right)^2 && (30)\\
&\leq \left(\int_B \frac{P(x)^2}{Q(x)}dx\right)\left(\int_B Q(x)dx\right) && (31)\\
&\leq \left(\int_{H^k} \frac{P(x)^2}{Q(x)}dx\right)\Pr_{D'}[B] \\
&= \mathrm{R}(D||D')\Pr_{D'}[B]. && (32)
\end{aligned}
$$

The derivation of (31) from (30) is made using the Cauchy-Scharwz inequality[21].

Proposition 11. *The minimum value of c such that $x^2/\sqrt{2x^2-1}$ is less than $1+c(x-1)^2$ for $x>1$ is $c=2$.*

Proof. We first show that for $x>1$, $x^2/\sqrt{2x^2-1} < 1+2(x-1)^2$. Let $p(x) = (2x^2-1)(1+2(x-1)^2)^2 - x^4$. Then $x^2/\sqrt{2x^2-1} < 1+2(x-1)^2$ if and only if $p(x)>0$. The polynomial $p(x)$ factors as $p(x) = (8x^3-8x^2+3x+9)(x-1)^3 = (8x^2(x-1)+3x+9)(x-1)^3$ which is a sum and product of positive numbers for $x>1$. Hence, it follows that $p(x)>0$ for $x>1$.

We next show that if the 2 in $1+2(x-1)^2$ is replaced by $c<2$, then the inequality $x^2/\sqrt{2x^2-1} < 1+c(x-1)^2$ cannot hold when x is close to 1. We set $\epsilon = x-1$ and $t = 4\epsilon + 2\epsilon^2$ and use the Taylor series $(1+t)^{-1/2} =$

[21] In the Cauchy-Scharwz inequality of the form $\left(\int_B f(x)g(x)dx\right)^2 \leq \left(\int_B f(x)^2dx\right)\left(\int_B g(x)^2dx\right)$, take $f(x) = P(x)/\sqrt{Q(x)}$ and $g(x) = \sqrt{Q(x)}$.

$1 - t/2 + 3t^2/8 \pm O(t^3)$. We have

$$\begin{aligned}\frac{x^2}{\sqrt{2x^2-1}} &= (1+\epsilon)^2(1+4\epsilon+2\epsilon^2)^{-1/2} \\ &= (1+2\epsilon+\epsilon^2)(1-2\epsilon-\epsilon^2+3(4\epsilon+2\epsilon^2)^2/8 \pm O(\epsilon^3)) \\ &= 1+2\epsilon^2 \pm O(\epsilon^3).\end{aligned}$$

For $c < 2$ and small ϵ, this is greater than $1 + c\epsilon^2$. □

Proposition 12. *Let $k \geq 1$ be a positive integer and $r_1, r_2 \in \mathbb{R}^+$ be such that $1 < r_2/r_1 < 1 + \sqrt{\ln(nk)/(nk)}/2$. Let D_{r_1} and D_{r_2} be the continuous Gaussian distributions on H having widths r_1 and r_2 respectively. Then*

$$\mathrm{R}(D^k_{r_1}||D^k_{r_2}) \leq \left(1 + \frac{1}{2}\cdot\frac{\ln(nk)}{nk}\right)^{nk}. \tag{33}$$

Proof. Direct calculation from the definition of the continuous Gaussian distribution D_r on H shows that for $r > 0$ and $x > 1/\sqrt{2}$, $\mathrm{R}(D_r||D_{xr}) = (x^2/\sqrt{2x^2-1})^n$. For $x > 1$, from Proposition 11, we have $(x^2/\sqrt{2x^2-1})^n$ is smaller than $(1 + 2(x-1)^2)^n$. So $\mathrm{R}(D^k_{r_1}||D^k_{r_2}) = (\mathrm{R}(D_{r_1}||D_{r_2}))^k \leq (1 + \ln(nk)/(2nk))^{nk}$. □

Setting $x = \alpha^2/\mathfrak{r}^2$ in the inequality $1 + x < (1 + x/2)^2$ for $x \neq 0$ and using (11), we have

$$1 \leq \frac{r'}{\mathfrak{r}} = \frac{\sqrt{r^2+\mathfrak{r}^2}}{\mathfrak{r}} \leq \sqrt{1+\frac{\alpha^2}{\mathfrak{r}^2}} < 1 + \frac{1}{2}\cdot\frac{\alpha^2}{\mathfrak{r}^2} = 1 + \frac{1}{2}\left(\frac{\ln(nN_2\ell)}{nN_2\ell}\right)^{1/2}. \tag{34}$$

Applying Proposition 12 with $k = N_2\ell$, $r_1 = \mathfrak{r}$ and $r_2 = r'$, we obtain

$$\mathrm{R}(D^{\ell N_2}_{\mathfrak{r}}||D^{\ell N_2}_{r'}) \leq (1 + \ln(nN_2\ell)/(2nN_2\ell))^{nN_2\ell}. \tag{35}$$

Proposition 13. *The right hand side of (35) is about $(nN_2\ell)^{1/2}$.*

Proof. The approximation can be seen by setting $x = (2nN_2\ell)/(\ln(nN_2\ell))$ and $m = (\ln(nN_2\ell))/2$ in $(1+1/x)^{mx} \approx e^m$. □

Remark 23. *If we follow the analysis in [21], then we obtain $(nN_2\ell)^3$ instead of $(nN_2\ell)^{1/2}$.*

Proposition 14. *For a good $s+t$, the measure of the set of good $\mathbf{z}$ under $D^{\ell N_2}_{r'}$ is at least about $\epsilon_2^2/(256nN_2\ell)^{1/2}$.*

Proof. From Section 4.4, we have that for a good $s + t$ the measure of the set of good $\mathbf{z}$ under $D^{\ell N_2}_{\mathfrak{r}}$ is at least $\epsilon_2/4$. In (32), considering B to be the set of good $\mathbf{z}$ and replacing k by $N_2\ell$, we have

$$\Pr_{D^{\ell N_2}_{r'}}[B] \geq \frac{\left(\Pr_{D^{\ell N_2}_{\mathfrak{r}}}[B]\right)^2}{\mathrm{R}(D^{\ell N_2}_{\mathfrak{r}}||D^{\ell N_2}_{r'})} \geq \frac{\epsilon_2^2}{16\mathrm{R}(D^{\ell N_2}_{\mathfrak{r}}||D^{\ell N_2}_{r'})} \gtrapprox \frac{\epsilon_2^2}{(256nN_2\ell)^{1/2}}.$$

□

Remark 24. *In [21], the ratio $\mathfrak{r}/\alpha$ is defined to be $((n\ell)/\ln(n\ell))^{1/4}$. If we use this definition of $\mathfrak{r}/\alpha$, and take $k = \ell$ in Proposition 12, then instead of (35) we would obtain*

$$\mathrm{R}(D^{\ell N_2}_{\mathfrak{r}}||D^{\ell N_2}_{r'}) \leq (\mathrm{R}(D^{\ell}_{\mathfrak{r}}||D^{\ell}_{r'}))^{N_2} \approx (n\ell)^{N_2/2}. \tag{36}$$

Since $N_2 > n^2$, this would lead to super-exponential running time $> n^{n^2}$.