Classical Hardness of Learning with Errors

Zvika Brakerski* Adeline Langlois † Chris Peikert‡ Oded Regev§ Damien Stehlé ¶

Abstract

We show that the Learning with Errors (LWE) problem is *classically* at least as hard as standard worst-case lattice problems, even with polynomial modulus. Previously this was only known under *quantum* reductions.

Our techniques capture the tradeoff between the dimension and the modulus of LWE instances, leading to a much better understanding of the landscape of the problem. The proof is inspired by techniques from several recent cryptographic constructions, most notably fully homomorphic encryption schemes.

1 Introduction

Over the last decade, lattices have emerged as a very attractive foundation for cryptography. The appeal of lattice-based primitives stems from the fact that their security can be based on *worst-case* hardness assumptions, that they appear to remain secure even against *quantum* computers, that they can be quite efficient, and that, somewhat surprisingly, for certain advanced tasks such as fully homomorphic encryption no other cryptographic assumption is known to suffice.

Virtually all recent lattice-based cryptographic schemes are based directly upon one of two natural average-case problems that have been shown to enjoy worst-case hardness guarantees: the *short integer solution* (SIS) problem and the *learning with errors* (LWE) problem. The former dates back to Ajtai's groundbreaking work [Ajt96], who showed that it is at least as hard as approximating several worst-case lattice problems, such as the (decision version of the) shortest vector problem, known as GapSVP, to within a polynomial factor in the lattice dimension. This hardness result was tightened in followup work (e.g., [MR04]), leading to a somewhat satisfactory understanding of the hardness of the SIS problem. The SIS problem has been the foundation for one-way [Ajt96] and collision-resistant hash functions [GGH96], identification schemes [MV03, Lyu08, KTX08], and digital signatures [GPV08, CHKP10, Boy10, MP12, Lyu12].

*Stanford University, zvika@stanford.edu. Supported by a Simons Postdoctoral Fellowship and DARPA.

†ENS de Lyon and Laboratoire LIP (U. Lyon, CNRS, ENS Lyon, INRIA, UCBL), 46 Allée d'Italie, 69364 Lyon Cedex 07, France. adeline.langlois@ens-lyon.fr.

‡School of Computer Science, College of Computing, Georgia Institute of Technology. This material is based upon work supported by the National Science Foundation under CAREER Award CCF-1054495, by DARPA under agreement number FA8750-11-C-0096, and by the Alfred P. Sloan Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, DARPA or the U.S. Government, or the Sloan Foundation. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

§Courant Institute, New York University. Supported by a European Research Council (ERC) Starting Grant. Part of the work done while the author was with the CNRS, DI, ENS, Paris.

¶ENS de Lyon and Laboratoire LIP (U. Lyon, CNRS, ENS Lyon, INRIA, UCBL), 46 Allée d'Italie, 69364 Lyon Cedex 07, France. damien.stehle@ens-lyon.fr. The author was partly supported by the Australian Research Council Discovery Grant DP110100628.

Our focus in this paper is on the latter problem, learning with errors. In this problem our goal is to distinguish with some non-negligible advantage between the following two distributions:

$$((\mathbf{a}_i, \langle \mathbf{a}_i, \mathbf{s} \rangle + e_i \bmod q))_i \quad \text{and} \quad ((\mathbf{a}_i, u_i))_i \,,$$

where $\mathbf{s}$ is chosen uniformly from $\mathbb{Z}_q^n$ and so are the $\mathbf{a}_i \in \mathbb{Z}_q^n$, u_i are chosen uniformly from $\mathbb{Z}_q$, and the "noise" $e_i \in \mathbb{Z}$ is sampled from some distribution supported on small numbers, typically a (discrete) Gaussian distribution with standard deviation αq for $\alpha = o(1)$.

The LWE problem has proved to be amazingly versatile, serving as the basis for a multitude of cryptographic constructions: secure public-key encryption under both chosen-plaintext [Reg05, PVW08, LP11] and chosen-ciphertext [PW08, Pei09, MP12] attacks, oblivious transfer [PVW08], identity-based encryption [GPV08, CHKP10, ABB10a, ABB10b], various forms of leakage-resilient cryptography (e.g., [AGV09, ACPS09, GKPV10]), fully homomorphic encryption [BV11, BGV12, Bra12] (following the seminal work of Gentry [Gen09]), and much more. It was also used to show hardness of learning problems [KS06].

Contrary to the SIS problem, however, the hardness of LWE is not sufficiently well understood. The main hardness reduction for LWE [Reg05] is similar to the one for SIS mentioned above, except that it is *quantum*. This means that the existence of an efficient algorithm for LWE, even a classical (i.e., non-quantum) one, only implies the existence of an efficient *quantum* algorithm for lattice problems. This state of affairs is quite unsatisfactory: even though one might conjecture that efficient quantum algorithms for lattice problems do not exist, our understanding of quantum algorithms is still at its infancy. It is therefore highly desirable to come up with a *classical* hardness reduction for LWE.

Progress in this direction was made by [Pei09] (with some simplifications in the followup by Lyubashevsky and Micciancio [LM09]). The main result there is that LWE with *exponential* modulus is as hard as some standard lattice problems using a classical reduction. As that hardness result crucially relies on the exponential modulus, the open question remained as to whether LWE is hard for smaller moduli, in particular polynomial moduli. In addition to being an interesting question in its own right, this question is of special importance since many cryptographic applications, as well as the learning theory result of Klivans and Sherstov [KS06], are instantiated in this setting. Some additional evidence that reducing the modulus is a fundamental question comes from the Learning Parity with Noise (LPN) problem, which can be seen as LWE with modulus 2 (albeit with a different error distribution), and whose hardness is a long-standing open question. We remark that [Pei09] does include a classical hardness of LWE with polynomial modulus, albeit one based on a non-standard lattice problem, whose hardness is arguably as debatable as that of the LWE problem itself.

To summarize, prior to our work, the existence of an efficient algorithm for LWE with polynomial modulus was only known to imply an efficient *quantum* algorithm for lattice problems, or an efficient classical algorithm for a non-standard lattice problem. While both consequences are unlikely, they are arguably not as earth-shattering as an efficient classical algorithm for lattice problems. Hence, some concern about the hardness of LWE persisted, tainting the plethora of cryptographic applications based on it.

Main result. We provide the first classical hardness reduction of LWE with polynomial modulus. Our reduction is the first to show that the existence of an efficient classical algorithm for LWE with any subexponential modulus would indeed have earth-shattering consequences: it would imply an efficient algorithm for worst-case instances of standard lattice problems.

Theorem 1.1 (Informal). *Solving n-dimensional* LWE *with $\mathrm{poly}(n)$ modulus implies an equally efficient solution to a worst-case lattice problem in dimension $\sqrt{n}$.*

As a result, we establish the hardness of all known applications of polynomial-modulus LWE based on classical worst-case lattice problems, previously only known under a quantum assumption.

Techniques. Even though our main theorem has the flavor of a statement in computational complexity, its proof crucially relies on a host of ideas coming from recent progress in cryptography, most notably recent breakthroughs in the construction of fully homomorphic encryption schemes.

At a high level, our main theorem is a "modulus reduction" result: we show a reduction from LWE with large modulus q and dimension n to LWE with (small) modulus $p = \mathrm{poly}(n)$ and dimension $n \log_2 q$. Theorem 1.1 now follows from the main result in [Pei09], which shows that the former problem with $q = 2^n$ is as hard as n-dimensional GapSVP. We note that the increase in dimension from n to $n \log_2 q$ is to be expected, as it essentially preserves the number of possible secrets (and hence the running time of the naive brute-force algorithm)..

Very roughly speaking, the main idea in modulus reduction is to map $\mathbb{Z}_q$ into $\mathbb{Z}_p$ through the naive mapping that sends any $a \in \{0, \ldots, q-1\}$ to $\lfloor pa/q \rfloor \in \{0, \ldots, p-1\}$. This basic idea is confounded by two issues. The first is that if carried out naively, this transformation introduces rounding artifacts into LWE, ruining the distribution of the output. We resolve this issue by using a more careful Gaussian randomized rounding procedure (Section 3). A second serious issue is that in order for the rounding errors not to be amplified when multiplied by the LWE secret $\mathbf{s}$, it is essential to assume that $\mathbf{s}$ has small coordinates. A major part of our reduction (Section 4) is therefore dedicated to showing a reduction from LWE (in dimension n) with arbitrary secret in $\mathbb{Z}_q^n$ to LWE (in dimension $n \log_2 q$) with a secret chosen uniformly over $\{0, 1\}$. This follows from a careful hybrid argument (Section 4.3) combined with a hardness reduction to the so-called "extended-LWE" problem, which is a variant of LWE in which we have some control over the error vector (Section 4.2).

We stress that even though our proof is inspired by and has analogues in the cryptographic literature, the details of the reductions are very different. In particular, the idea of modulus reduction plays a key role in recent work on fully homomorphic encryption schemes, giving a way to control the noise growth during homomorphic operations [BV11, BGV12, Bra12]. However, since the goal there is merely to preserve the functionality of the scheme, their modulus reduction can be performed in a rather naive way similar to the one outlined above, and so the output of their procedure does not constitute a valid LWE instance. In our reduction we need to perform a much more delicate modulus reduction, which we do using Gaussian randomized rounding, as mentioned above.

The idea of reducing LWE to have a $\{0, 1\}$ secret also exists already in the cryptographic literature: precisely such a reduction was shown by Goldwasser et al. [GKPV10] who were motivated by questions in leakage-resilient cryptography. Their reduction, however, incurred a severe blow-up in the noise rate, making it useless for our purposes. In more detail, not being able to faithfully reproduce the LWE distribution in the output, they resort to hiding the faults in the output distribution under a huge independent fresh noise, in order to make it close to the correct one. The trouble with this "noise flooding" approach is that the amount of noise one has to add depends on the running time of the algorithm solving the target $\{0, 1\}$-LWE problem, which in turn forces the modulus to be equally big. So while in principle we could use the reduction from [GKPV10] (and shorten our proof by about a half), this would lead to a qualitatively much weaker result: the modulus and the approximation ratio for the worst-case lattice problem would both grow with the running time of the $\{0, 1\}$-LWE algorithm. In particular, we would not be able to show that for some fixed polynomial modulus, LWE is a hard problem; instead, in order to capture all polynomial time algorithms, we would have to take a super-polynomial modulus, and rely on the hardness of worst-case lattice problem to within super-polynomial approximation factors. In contrast, with our reduction, the modulus and the

approximation ratio both remain fixed independently of the target $\{0,1\}$-LWE algorithm.

As mentioned above, our alternative to the reduction in [GKPV10] is based on a hybrid argument combined with a new hardness reduction for the "extended LWE" problem, which is a variant of LWE in which in addition to the LWE samples, we also get to see the inner product of the vector of error terms with a vector $\mathbf{z}$ of our choosing. This problem has its origins in the cryptographic literature, namely in the work of O'Neill, Peikert, and Waters [OPW11] on (bi)deniable encryption and the later work of Alperin-Sheriff and Peikert [AP12] on key-dependent message security. The hardness reductions included in those papers are not sufficient for our purposes, as they cannot handle large moduli or error terms, which is crucial in our setting. We therefore provide an alternative reduction which is conceptually much simpler, and essentially subsumes both previous reductions. Our reduction works equally well with exponential moduli and correspondingly long error vectors, a case earlier reductions could not handle.

Broader perspective. As a byproduct of the proof of Theorem 1.1, we obtain several results that shed new light on the hardness of LWE. Most notably, our modulus reduction result in Section 3 is actually far more general, and can be used to show a "modulus expansion/dimension reduction" tradeoff. Namely, it shows a reduction from LWE in dimension n and modulus p to LWE in dimension n/k and modulus p^k (see Corollary 3.4). Combined with our modulus reduction, this has the following interesting consequence: the hardness of n-dimensional LWE with modulus q is a function of the quantity $n \log_2 q$. In other words, varying n and q individually while keeping $n \log_2 q$ fixed essentially preserves the hardness of LWE.

Although we find this statement quite natural (since $n \log_2 q$ represents the number of bits in the secret), it has some surprising consequences. One is that n-dimensional LWE with modulus 2^n is essentially as hard as n^2-dimensional LWE with polynomial modulus. As a result, n-dimensional LWE with modulus 2^n, which was shown in [Pei09] to be as hard as n-dimensional lattice problems using a classical reduction, is actually as hard as n^2-dimensional lattice problems using a quantum reduction. The latter is presumably a much harder problem, requiring $\exp(\widetilde{\Omega}(n^2))$ time to solve. This corollary highlights an inherent quadratic loss in the classical reduction of [Pei09] (and as a result also our Theorem 1.1) compared to the quantum one in [Reg05].

A second interesting consequence is that 1-dimensional LWE with modulus 2^n is essentially as hard as n-dimensional LWE with polynomial modulus. The 1-dimensional version of LWE is closely related to the Hidden Number Problem of Boneh and Venkatesan [BV96]. It is also essentially equivalent to the Ajtai-Dwork-type [AD97] cryptosystem in [Reg03], as follows from simple reductions similar to the one in the appendix of [Reg10a]. Moreover, the 1-dimensional version can be seen as a special case of the Ring-LWE problem introduced in [LPR10] (for ring dimension 1, i.e., ring equal to $\mathbb{Z}$). This allows us, via the ring switching technique from [GHPS12], to obtain the first hardness proof of Ring-LWE, with arbitrary ring dimension and exponential modulus, under the hardness of problems on general lattices (as opposed to just ideal lattice problems). In addition, this leads to the first hardness proof for the Ring-SIS problem [LM06, PR06] with exponential modulus under the hardness of general lattice problems, via the standard LWE-to-SIS reduction. (We note that since both results are obtained by scaling up from a ring of dimension 1, the hardness does not improve as the ring dimension increases.)

A final interesting consequence of our reductions is that (the decision form of) LWE is hard with an arbitrary huge modulus, e.g., a prime; see Corollary 3.3. Previous results (e.g., [Reg05, Pei09, MM11, MP12]) required the modulus to be *smooth*, i.e., all its prime divisors had to be polynomially bounded.

Open questions. As mentioned above, our Theorem 1.1 inherits from [Pei09] a quadratic loss in the dimension, which does not exist in the quantum reduction [Reg05] nor in the known hardness reductions

for SIS. At a technical level, this quadratic loss stems from the fact that the reduction in [Pei09] is not iterative. In contrast, the quantum reduction in [Reg05] as well as the reductions for SIS are iterative, and as a result do not incur the quadratic loss. We note that an additional side effect of the non-iterative reduction is that the hardness in Theorem 1.1 and [Pei09] is based only on the worst-case lattice problem GapSVP (and the essentially equivalent BDD and uSVP [LM09]), and not on problems like SIVP, which the quantum reduction of [Reg05] and the hardness reductions for SIS can handle. One case where this is very significant is when dealing with ideal lattices, as in the hardness reduction for Ring-LWE, since GapSVP turns out to be an easy problem there.

We therefore believe that it is important to understand whether there exists a classical reduction that does not incur the quadratic loss inherent in [Pei09] and in Theorem 1.1. In other words, is n-dimensional LWE with polynomial modulus classically as hard as n-dimensional lattice problems (as opposed to $\sqrt{n}$-dimensional)? This would constitute the first full dequantization of the quantum reduction in [Reg05].

While it is natural to conjecture that the answer to this question is positive, a negative answer would be quite tantalizing. In particular, it is conceivable that there exists a (classical) algorithm for LWE with polynomial modulus running in time $2^{O(\sqrt{n})}$. Due to the quadratic expansion in Theorem 1.1, this would not lead to a faster classical algorithm for lattice problems; it would, however, lead to a $2^{O(\sqrt{n})}$-time *quantum* algorithm for lattice problems using the reduction in [Reg05]. The latter would be a major progress in quantum algorithms, yet is not entirely unreasonable; in fact, a $2^{O(\sqrt{n})}$-time quantum algorithm for a somewhat related quantum task was discovered by Kuperberg [Kup05] (see also [Reg02]).

2 Preliminaries

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ denote the cycle, i.e., the additive group of reals modulo 1. We also denote by $\mathbb{T}_q$ its cyclic subgroup of order q, i.e., the subgroup given by $\{0, 1/q, \ldots, (q-1)/q\}$.

For two probability distributions P, Q over some discrete domain, we define their statistical distance as $\sum |P(i) - Q(i)|/2$ where i ranges over the distribution domain, and extend this to continuous distributions in the obvious way. We recall the following easy fact (see, e.g., [AD87, Eq. (2.3)] for a proof).

Claim 2.1. *If P and Q are two probability distributions such that $P(i) \geq (1-\varepsilon)Q(i)$ holds for all i, then the statistical distance between P and Q is at most ε.*

We will use the following immediate corollary of the leftover hash lemma [HILL99].

Lemma 2.2. *Let $k, n, q \geq 1$ be integers, and $\varepsilon > 0$ be such that $n \geq k \log_2 q + 2\log_2(1/\varepsilon)$. For $\mathbf{H} \leftarrow \mathbb{T}_q^{k \times n}$, $\mathbf{z} \leftarrow \{0,1\}^n$, $\mathbf{u} \leftarrow \mathbb{T}_q^k$, the distributions of $(\mathbf{H}, \mathbf{Hz})$ and $(\mathbf{H}, \mathbf{u})$ are within statistical distance at most ε.*

A *distinguishing problem* P is defined by two distributions P_0 and P_1, and a solution to the problem is the ability to distinguish between these distributions. The *advantage* of an algorithm $\mathcal{A}$ with binary output on P is defined as

$$\mathrm{Adv}[\mathcal{A}] = \left|\Pr[\mathcal{A}(P_0)] - \Pr[\mathcal{A}(P_1)]\right| .$$

A reduction from a problem P to a problem Q is an efficient (i.e., polynomial-time) algorithm $\mathcal{A}^{\mathcal{B}}$ that solves P given access to an oracle $\mathcal{B}$ that solves Q. Most of our reductions (in fact all except the one in Lemma 2.13) are what we call "transformation reductions:" these reductions perform some transformation to the input and then apply the oracle to the result.

2.1 Lattices

An n-dimensional (full-rank) lattice $\Lambda \subseteq \mathbb{R}^n$ is the set of all integer linear combinations of some set of n linearly independent *basis* vectors $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\} \subseteq \mathbb{R}^n$,

$$\Lambda = \mathcal{L}(\mathbf{B}) = \Big\{ \sum_{i \in [n]} z_i \mathbf{b}_i \ : \ \mathbf{z} \in \mathbb{Z}^n \Big\}.$$

The *dual lattice* of $\Lambda \subset \mathbb{R}^n$ is defined as $\Lambda^* = \{\mathbf{x} \in \mathbb{R}^n : \langle \Lambda, \mathbf{x} \rangle \subseteq \mathbb{Z}\}$.

The *minimum distance* (or *first successive minimum*) $\lambda_1(\Lambda)$ of a lattice Λ is the length of a shortest nonzero lattice vector, i.e., $\lambda_1(\Lambda) = \min_{\mathbf{0} \neq \mathbf{x} \in \Lambda} \|\mathbf{x}\|$. For an approximation ratio $\gamma = \gamma(n) \geq 1$, the GapSVP_γ is the problem of deciding, given a basis $\mathbf{B}$ of an n-dimensional lattice $\Lambda = \mathcal{L}(\mathbf{B})$ and a number d, between the case where $\lambda_1(\mathcal{L}(\mathbf{B})) \leq d$ and the case where $\lambda_1(\mathcal{L}(\mathbf{B})) > \gamma d$. We refer to [Kho10, Reg10b] for a recent account on the computational complexity of GapSVP_γ.

2.2 Gaussian measures

For $r > 0$, the n-dimensional Gaussian function $\rho_r : \mathbb{R}^n \to (0, 1]$ is defined as

$$\rho_r(\mathbf{x}) := \exp(-\pi \|\mathbf{x}\|^2 / r^2).$$

We extend this definition to sets, i.e., $\rho_r(A) = \sum_{\mathbf{x} \in A} \rho_r(\mathbf{x}) \in [0, +\infty]$ for any $A \subseteq \mathbb{R}^n$. The (spherical) continuous Gaussian distribution D_r is the distribution with density function proportional to ρ_r. More generally, for a matrix $\mathbf{B}$, we denote by $D_{\mathbf{B}}$ the distribution of $\mathbf{Bx}$ where $\mathbf{x}$ is sampled from D_1. When $\mathbf{B}$ is nonsingular, its probability density function is proportional to

$$\exp(-\pi \mathbf{x}^T (\mathbf{B}\mathbf{B}^T)^{-1} \mathbf{x}).$$

A basic fact is that for any matrices $\mathbf{B}_1, \mathbf{B}_2$, the sum of a sample from $D_{\mathbf{B}_1}$ and an independent sample from $D_{\mathbf{B}_2}$ is distributed like $D_{\mathbf{C}}$ for $\mathbf{C} = (\mathbf{B}_1 \mathbf{B}_1^T + \mathbf{B}_2 \mathbf{B}_2^T)^{1/2}$.

For an n-dimensional lattice Λ and a vector $\mathbf{u} \in \mathbb{R}^n$, we define the *discrete Gaussian distribution* $D_{\Lambda + \mathbf{u}, r}$ as the discrete distribution with support on the coset $\Lambda + \mathbf{u}$ whose probability mass function is proportional to ρ_r. There exists an efficient procedure that samples within negligible statistical distance of any (not too narrow) discrete Gaussian distribution ([GPV08, Theorem 4.1]; see also [Pei10]). In the next lemma, proved in Section 5, we modify this sampler so that the output is distributed exactly as a discrete Gaussian. This also allows us to sample from slightly narrower Gaussians. Strictly speaking, the lemma is not needed for our results, and we could use instead the original sampler from [GPV08]. Using our exact sampler leads to slightly cleaner proofs as well as a (miniscule) improvement in the parameters of our reductions, and we include it here mainly in the hope that it finds further applications in the future.

Lemma 2.3. *There is a probabilistic polynomial-time algorithm that, given a basis $\mathbf{B}$ of an n-dimensional lattice $\Lambda = \mathcal{L}(\mathbf{B})$, $\mathbf{c} \in \mathbb{R}^n$, and a parameter $r \geq \|\widetilde{\mathbf{B}}\| \cdot \sqrt{\ln(2n + 4)/\pi}$, outputs a sample distributed according to $D_{\Lambda + \mathbf{c}, r}$.*

Here, $\widetilde{\mathbf{B}}$ denotes the Gram-Schmidt orthogonalization of $\mathbf{B}$, and $\|\widetilde{\mathbf{B}}\|$ is the length of the longest vector in it. We recall the definition of the *smoothing parameter* from [MR04].

Definition 2.4. *For a lattice Λ and positive real $\varepsilon > 0$, the smoothing parameter $\eta_\varepsilon(\Lambda)$ is the smallest real $s > 0$ such that $\rho_{1/s}(\Lambda^* \setminus \{\mathbf{0}\}) \leq \varepsilon$.*

Lemma 2.5 ([GPV08, Lemma 3.1]). *For any $\varepsilon > 0$ and n-dimensional lattice Λ with basis $\mathbf{B}$,*

$$\eta_\varepsilon(\Lambda) \leq \|\widetilde{\mathbf{B}}\| \sqrt{\ln(2n(1+1/\varepsilon))/\pi}.$$

We now collect some known facts on Gaussian distributions and lattices.

Lemma 2.6 ([MR04, Lemma 4.1]). *For any n-dimensional lattice Λ, $\varepsilon > 0$, $r \geq \eta_\varepsilon(\Lambda)$, the distribution of $\mathbf{x} \bmod \Lambda$ where $\mathbf{x} \leftarrow D_r$ is within statistical distance $\varepsilon/2$ of the uniform distribution on cosets of Λ.*

Lemma 2.7 ([Reg05, Claim 3.8]). *For any n-dimensional lattice Λ, $\varepsilon > 0$, $r \geq \eta_\varepsilon(\Lambda)$, and $\mathbf{c} \in \mathbb{R}^n$, we have $\rho_r(\Lambda + \mathbf{c}) \in [\frac{1-\varepsilon}{1+\varepsilon}, 1] \cdot \rho_r(\Lambda)$.*

Lemma 2.8 ([Reg05, Claim 3.9]). *Let Λ be an n-dimensional lattice, let $\mathbf{u} \in \mathbb{R}^n$ be arbitrary, let $r, s > 0$ and let $t = \sqrt{r^2 + s^2}$. Assume that $rs/t = 1/\sqrt{1/r^2 + 1/s^2} \geq \eta_\varepsilon(\Lambda)$ for some $\varepsilon < 1/2$. Consider the continuous distribution Y on $\mathbb{R}^n$ obtained by sampling from $D_{\Lambda+\mathbf{u},r}$ and then adding a noise vector taken from D_s. Then, the statistical distance between Y and D_t is at most 4ε.*

Lemma 2.9 ([Reg05, Corollary 3.10]). *Let Λ be an n-dimensional lattice, let $\mathbf{u}, \mathbf{z} \in \mathbb{R}^n$ be arbitrary, and let $r, \alpha > 0$. Assume that $(1/r^2 + (\|\mathbf{z}\|/\alpha)^2)^{-1/2} \geq \eta_\varepsilon(\Lambda)$ for some $\varepsilon < 1/2$. Then the distribution of $\langle \mathbf{z}, \mathbf{v} \rangle + e$ where $\mathbf{v} \leftarrow D_{\Lambda+\mathbf{u},r}$ and $e \leftarrow D_\alpha$, is within statistical distance 4ε of D_β for $\beta = \sqrt{(r\|\mathbf{z}\|)^2 + \alpha^2}$.*

Lemma 2.10 (Special case of [Pei10, Theorem 3.1]). *Let Λ be a lattice and $r, s > 0$ be such that $s \geq \eta_\varepsilon(\Lambda)$ for some $\varepsilon \leq 1/2$. Then if we choose $\mathbf{x}$ from the continuous Gaussian D_r and then choose $\mathbf{y}$ from the discrete Gaussian $D_{\Lambda-\mathbf{x},s}$ then $\mathbf{x} + \mathbf{y}$ is within statistical distance 8ε of the discrete Gaussian $D_{\Lambda,(r^2+s^2)^{1/2}}$.*

2.3 Learning with Errors

For integers $n, q \geq 1$, an integer vector $\mathbf{s} \in \mathbb{Z}^n$, and a probability distribution ϕ on $\mathbb{R}$, let $A_{q,\mathbf{s},\phi}$ be the distribution over $\mathbb{T}_q^n \times \mathbb{T}$ obtained by choosing $\mathbf{a} \in \mathbb{T}_q^n$ uniformly at random and an error term e from ϕ, and outputting the pair $(\mathbf{a}, b = \langle \mathbf{a}, \mathbf{s} \rangle + e) \in \mathbb{T}_q^n \times \mathbb{T}$.

Definition 2.11. *For integers $n, q \geq 1$, an error distribution ϕ over $\mathbb{R}$, and a distribution $\mathcal{D}$ over $\mathbb{Z}^n$, the (average-case) decision variant of the* LWE *problem, denoted* $\mathsf{LWE}_{n,q,\phi}(\mathcal{D})$*, is to distinguish given arbitrarily many independent samples, the uniform distribution over $\mathbb{T}_q^n \times \mathbb{T}$ from $A_{q,\mathbf{s},\phi}$ for a fixed $\mathbf{s}$ sampled from $\mathcal{D}$. The variant where the algorithm only gets a bounded number of samples $m \in \mathbb{N}$ is denoted* $\mathsf{LWE}_{n,m,q,\phi}(\mathcal{D})$.

Notice that the distribution $A_{q,\mathbf{s},\phi}$ only depends on $\mathbf{s} \bmod q$, and so one can assume without loss of generality that $\mathbf{s} \in \{0, \ldots, q-1\}^n$. Moreover, using a standard random self-reduction, for any distribution over secrets $\mathcal{D}$, one can reduce $\mathsf{LWE}_{n,q,\phi}(\mathcal{D})$ to $\mathsf{LWE}_{n,q,\phi}(U(\{0, \ldots, q-1\}^n))$, and we will occasionally use $\mathsf{LWE}_{n,q,\phi}$ to denote the latter (as is common in previous work). When the noise is a Gaussian with parameter $\alpha > 0$, i.e., $\phi = D_\alpha$, we use the shorthand $\mathsf{LWE}_{n,q,\alpha}(\mathcal{D})$. We note that by discretizing the error using Lemma 2.10 and using the so-called "normal form" of LWE (see [ACPS09]), one can efficiently reduce $\mathsf{LWE}_{n,q,\alpha}$ to $\mathsf{LWE}_{n,q,\alpha}(\mathcal{D})$ where $\mathcal{D}$ is the discrete Gaussian distribution $D_{\mathbb{Z}^n,\sqrt{2}\alpha q}$, as long as $\alpha q \geq \sqrt{n}$. Finally, since the case when $\mathcal{D}$ is uniform over $\{0,1\}^n$ plays an important role in this paper, we will denote it by $\mathsf{binLWE}_{n,q,\phi}$ (and by $\mathsf{binLWE}_{n,m,q,\phi}$ when the algorithm only gets m samples).

Unknown (Bounded) Noise Rate. We also consider a variant of LWE in which the amount of noise is some unknown $\beta \leq \alpha$ (as opposed to exactly α), with β possibly depending on the secret $\mathbf{s}$. As the following lemma shows, this does not make the problem significantly harder.

Definition 2.12. *For integers $n, q \geq 1$ and $\alpha \in (0,1)$, $\mathsf{LWE}_{n,q,\leq\alpha}$ is the problem of solving $\mathsf{LWE}_{n,q,\beta}$ for any $\beta = \beta(\mathbf{s}) \leq \alpha$.*

Lemma 2.13. *Let $\mathcal{A}$ be an algorithm for $\mathsf{LWE}_{n,m,q,\alpha}$ with advantage at least $\varepsilon > 0$. Then there exists an algorithm $\mathcal{B}$ for $\mathsf{LWE}_{n,m',q,\leq\alpha}$ using oracle access to $\mathcal{A}$ and with advantage $\geq 1/3$, where both m' and its running time are $\mathrm{poly}(m, 1/\varepsilon, n, \log q)$.*

The proof is standard (see, e.g., [Reg05, Lemma 3.7] for the analogous statement for the search version of LWE). The idea is to use Chernoff bound to estimate $\mathcal{A}$'s success probability on the uniform distribution, and then add noise in small increments to our given distribution and estimate $\mathcal{A}$'s behavior on the resulting distributions. If there is a gap between any of these and the uniform behavior, the input distribution is deemed non-uniform. The full proof is omitted.

Relation to Lattice Problems. Regev [Reg05] and Peikert [Pei09] showed quantum and classical reductions (respectively) from the worst-case hardness of the GapSVP problem to the search version of LWE. (We note that the quantum reduction in [Reg05] also shows a reduction from SIVP.) As mentioned in the introduction, the classical reduction only works when the modulus q is exponential in the dimension n. This is summarized in the following theorem, which is derived from [Reg05, Theorem 3.1] and [Pei09, Theorem 3.1].

Theorem 2.14. *Let $n, q \geq 1$ be integers and let $\alpha \in (0,1)$ be such that $\alpha q \geq 2\sqrt{n}$. Then there exists a quantum reduction from worst-case n-dimensional $\mathsf{GapSVP}_{\widetilde{O}(n/\alpha)}$ to $\mathsf{LWE}_{n,q,\alpha}$. If in addition $q \geq 2^{n/2}$ then there is also a classical reduction between those problems.*

In order to obtain hardness of the *decision* version of LWE, which is the one we consider throughout the paper, one employs a search-to-decision reduction. Several such reductions appear in the literature (e.g., [Reg05, Pei09, MP12]). The most recent reduction by Micciancio and Peikert [MP12], which essentially subsumes all previous reductions, requires the modulus q to be smooth. Below we give the special case when the modulus is a power of 2, which suffices for our purposes. It follows from our results that (decision) LWE is hard not just for a smooth modulus q, as follows from [MP12], but actually for all moduli q, including prime moduli, with only a small deterioration in the noise (see Corollaries 3.2 and 3.3).

Theorem 2.15 (Special case of [MP12, Theorem 3.1]). *Let q be a power of 2, and α satisfy $1/q < \alpha < 1/\omega(\sqrt{\log n})$. Then there exists an efficient reduction from search $\mathsf{LWE}_{n,q,\alpha}$ to (decision) $\mathsf{LWE}_{n,q,\alpha'}$ for $\alpha' = \alpha \cdot \omega(\log n)$.*

3 Modulus-Dimension Switching

The main results of this section are Corollaries 3.2 and 3.4 below. Both are special cases of the following technical theorem. We say that a distribution $\mathcal{D}$ over $\mathbb{Z}^n$ is (B, δ)-bounded for some reals $B, \delta \geq 0$ if the probability that $\mathbf{x} \leftarrow \mathcal{D}$ has norm greater than B is at most δ.

Theorem 3.1. *Let $m, n, n', q, q' \geq 1$ be integers, let $\mathbf{G} \in \mathbb{Z}^{n' \times n}$ be such that the lattice $\Lambda = \frac{1}{q'}\mathbf{G}^T\mathbb{Z}^{n'} + \mathbb{Z}^n$ has a known basis $\mathbf{B}$, and let $\mathcal{D}$ be an arbitrary (B, δ)-bounded distribution over $\mathbb{Z}^n$. Let $\alpha, \beta > 0$ and $\varepsilon \in (0, 1/2)$ satisfy*

$$\beta^2 \geq \alpha^2 + (4/\pi)\ln(2n(1+1/\varepsilon)) \cdot (\max\{q^{-1}, \|\widetilde{\mathbf{B}}\|\} \cdot B)^2.$$

Then there is an efficient (transformation) reduction from $\mathsf{LWE}_{n,m,q,\leq\alpha}(\mathcal{D})$ to $\mathsf{LWE}_{n',m,q',\leq\beta}(\mathbf{G} \cdot \mathcal{D})$ that reduces the advantage by at most $\delta + 14\varepsilon m$.

Here we use the notation $\|\widetilde{\mathbf{B}}\|$ from Lemma 2.3. We also note that if needed, the distribution on secrets produced by the reduction can always be turned into the uniform distribution on $\mathbb{Z}_{q'}^{n'}$, as mentioned after Definition 2.11. Also, we recall that there exists an elementary reduction from $\mathsf{LWE}_{n',q',\leq\beta}$ to $\mathsf{LWE}_{n',q',\beta}$ (see Lemma 2.13).

Here we state two important corollaries of the theorem. The first corresponds to just modulus reduction (the LWE dimension is preserved), and is obtained by letting $n' = n$, $\mathbf{G} = \mathbf{I}$ be the n-dimensional identity matrix, and $\mathbf{B} = \mathbf{I}/q'$. For example, we can take $q \geq q' \geq \sqrt{2\ln(2n(1+1/\varepsilon))} \cdot (B/\alpha)$ and $\beta = \sqrt{2}\alpha$, which corresponds to reducing an arbitrary modulus to almost B/α, while increasing the initial error rate α by just a small constant factor.

Corollary 3.2. *For any $m, n \geq 1$, $q \geq q' \geq 1$, (B, δ)-bounded distribution $\mathcal{D}$ over $\mathbb{Z}^n$, $\alpha, \beta > 0$ and $\varepsilon \in (0, 1/2)$ such that*

$$\beta^2 \geq \alpha^2 + (4/\pi)\ln(2n(1+1/\varepsilon)) \cdot (B/q')^2,$$

there is an efficient reduction from $\mathsf{LWE}_{n,m,q,\leq\alpha}(\mathcal{D})$ to $\mathsf{LWE}_{n,m,q',\leq\beta}(\mathcal{D})$ that reduces the advantage by at most $\delta + 14\varepsilon m$.

In particular, by using the normal form of LWE, in which the secret has distribution $\mathcal{D} = D_{\mathbb{Z}^n,\alpha q}$, we can switch to a power-of-2 modulus with only a small loss in the noise rate, as described in the following corollary. Together with the known search-to-decision reduction (Theorem 2.15), this extends the known hardness of (decision) LWE to *any* modulus q. Here we use that $\mathcal{D} = D_{\mathbb{Z}^n,r}$ is $(Cr\sqrt{n\log(n/\delta)}, \delta)$-bounded for some universal constant $C > 0$, which follows by taking union bound over the n coordinates. (Alternatively, one could use that it is $(r\sqrt{n}, 2^{-n})$-bounded, as follows from [Ban93, Lemma 1.5], leading to a slightly tighter statement for large n.)

Corollary 3.3. *Let $\alpha > 0$, $\delta \in (0, 1/2)$, $m \geq n \geq 1$, $q' \geq 1$, and let $q \in [q', 2q')$ be the smallest power of 2 not smaller than q'. There exists an efficient (transformation) reduction from $\mathsf{LWE}_{n,m,q,\leq\alpha}$ to $\mathsf{LWE}_{n,m,q',\leq\beta}$ where*

$$\beta = C\alpha\sqrt{n}\sqrt{\log(n/\delta)\log(nm/\delta)}$$

for some universal constant $C > 0$, losing at most δ in the advantage.

Another corollary illustrates a modulus-dimension tradeoff. Assume $n = kn'$ for some $k \geq 1$, and let $q' = q^k$. Let $\mathbf{G} = \mathbf{I}_{n'} \otimes \mathbf{g}$, where $\mathbf{g} = (1, q, q^2, \ldots, q^{k-1})^T \in \mathbb{Z}^k$. We then have $\Lambda = q^{-k}\mathbf{G}^T\mathbb{Z}^{n'} + \mathbb{Z}^n$. A basis of Λ is given by

$$\mathbf{B} = \mathbf{I}_{n'} \otimes \begin{bmatrix} q^{-1} & q^{-2} & \cdots & q^{-k} \\ & q^{-1} & \cdots & q^{1-k} \\ & & \ddots & \vdots \\ & & & q^{-1} \end{bmatrix} \in \mathbb{R}^{n \times n};$$

this is since the column vectors of $\mathbf{B}$ belong to Λ and the determinants match. Orthogonalizing from left to right, we have $\widetilde{\mathbf{B}} = q^{-1}\mathbf{I}$ and so $\|\widetilde{\mathbf{B}}\| = q^{-1}$. We therefore obtain the following corollary, showing that we can trade off the dimension against the modulus, holding $n \log q = n' \log q'$ fixed. For example, letting $\mathcal{D} = D_{\mathbb{Z}^n, \alpha q}$ (corresponding to a secret in normal form), which is $(\alpha q \sqrt{n}, 2^{-n})$-bounded, the reduction increases the error rate by about a $\sqrt{n}$ factor.

Corollary 3.4. *For any $n, m, q \geq 1$, $k \geq 1$ that divides n, (B, δ)-bounded distribution $\mathcal{D}$ over $\mathbb{Z}^n$, $\alpha, \beta > 0$, and $\varepsilon \in (0, 1/2)$ such that*

$$\beta^2 \geq \alpha^2 + (4/\pi) \ln(2n(1 + 1/\varepsilon)) \cdot (B/q)^2,$$

there is an efficient reduction from $\mathsf{LWE}_{n,m,q,\leq\alpha}(\mathcal{D})$ to $\mathsf{LWE}_{n/k,m,q^k,\leq\beta}(\mathbf{G} \cdot \mathcal{D})$ that reduces the advantage by at most $\delta + 14\varepsilon m$, where $\mathbf{G} = \mathbf{I}_{n/k} \otimes (1, q, q^2, \ldots, q^{k-1})^T$.

Theorem 3.1 follows immediately from the following lemma.

Lemma 3.5. *Adopt the notation of Theorem 3.1, and let*

$$r \geq \max\{q^{-1}, \|\widetilde{\mathbf{B}}\|\} \cdot \sqrt{2 \ln(2n(1 + 1/\varepsilon))/\pi}.$$

There is an efficient mapping from $\mathbb{T}_q^n \times \mathbb{T}$ to $\mathbb{T}_{q'}^{n'} \times \mathbb{T}$, which has the following properties:

- *If the input is uniformly random, then the output is within statistical distance 4ε from the uniform distribution.*
- *If the input is distributed according to $A_{q,\mathbf{s},D_\alpha}$ for some $\mathbf{s} \in \mathbb{Z}^n$ with $\|\mathbf{s}\| \leq B$, then the output distribution is within statistical distance 10ε from $A_{q',\mathbf{G}\mathbf{s},D_{\alpha'}}$, where $(\alpha')^2 = \alpha^2 + r^2(\|\mathbf{s}\|^2 + B^2) \leq \alpha^2 + 2(rB)^2$.*

Proof. The main idea behind the reduction is to encode $\mathbb{T}_q^n$ into $\mathbb{T}_{q'}^{n'}$, so that the mod-1 inner products between vectors in $\mathbb{T}_q^n$ and a short vector $\mathbf{s} \in \mathbb{Z}^n$, and between vectors in $\mathbb{T}_{q'}^{n'}$ and $\mathbf{G}\mathbf{s} \in \mathbb{Z}^{n'}$, are nearly equivalent. In a bit more detail, the reduction will map its input vector $\mathbf{a} \in \mathbb{T}_q^n$ (from the given LWE-or-uniform distribution) to a vector $\mathbf{a}' \in \mathbb{T}_{q'}^{n'}$, so that

$$\langle \mathbf{a}', \mathbf{G}\mathbf{s} \rangle = \langle \mathbf{G}^T \mathbf{a}', \mathbf{s} \rangle \approx \langle \mathbf{a}, \mathbf{s} \rangle \bmod 1$$

for any (unknown) $\mathbf{s} \in \mathbb{Z}^n$. To do this, it randomly samples $\mathbf{a}'$ so that $\mathbf{G}^T \mathbf{a}' \approx \mathbf{a} \bmod \mathbb{Z}^n$, where the approximation error will be a discrete Gaussian of parameter r.

We can now formally define the reduction, which works as follows. On an input pair $(\mathbf{a}, b) \in \mathbb{T}_q^n \times \mathbb{T}$, it does the following:

- Choose $\mathbf{f} \leftarrow D_{\Lambda - \mathbf{a}, r}$ using Lemma 2.3 with basis $\mathbf{B}$, and let $\mathbf{v} = \mathbf{a} + \mathbf{f} \in \Lambda/\mathbb{Z}^n$. (The coset $\Lambda - \mathbf{a}$ is well defined since $\mathbf{a} = \bar{\mathbf{a}} + \mathbb{Z}^n$ is some coset of $\mathbb{Z}^n \subseteq \Lambda$.) Choose a uniformly random solution $\mathbf{a}' \in \mathbb{T}_{q'}^{n'}$ to the equation $\mathbf{G}^T \mathbf{a}' = \mathbf{v} \bmod \mathbb{Z}^n$. This can be done by computing a basis of the solution set $\mathbf{G}^T \mathbf{a}' = \mathbf{0} \bmod \mathbb{Z}^n$, and adding a uniform element from that set to an arbitrary solution to the equation $\mathbf{G}^T \mathbf{a}' = \mathbf{v} \bmod \mathbb{Z}^n$.
- Choose $e' \leftarrow D_{rB}$ and let $b' = b + e' \in \mathbb{T}$.
- Output $(\mathbf{a}', b')$.

We now analyze the reduction. First, if the distribution of the input is uniform, then it suffices to show that $\mathbf{a}'$ is (nearly) uniformly random, because both b and e' are independent of $\mathbf{a}'$, and $b \in \mathbb{T}$ is uniform. To prove this claim, notice that it suffices to show that the coset $\mathbf{v} \in \Lambda/\mathbb{Z}^n$ is (nearly) uniformly random, because each $\mathbf{v}$ has the same number of solutions $\mathbf{a}'$ to $\mathbf{G}^T\mathbf{a}' = \mathbf{v} \bmod \mathbb{Z}^n$. Next, observe that for any $\bar{\mathbf{a}} \in \mathbb{T}_q^n$ and $\bar{\mathbf{f}} \in \Lambda - \bar{\mathbf{a}}$, we have by Lemma 2.7 (using that $r \geq \eta_\varepsilon(\Lambda)$ by Lemma 2.5) that

$$\begin{aligned}\Pr[\mathbf{a} = \bar{\mathbf{a}} \wedge \mathbf{f} = \bar{\mathbf{f}}] &= q^{-n} \cdot \rho_r(\bar{\mathbf{f}})/\rho_r(\Lambda - \bar{\mathbf{a}}) \\ &\in C\big[1, \tfrac{1+\varepsilon}{1-\varepsilon}\big] \cdot \rho_r(\bar{\mathbf{f}}). \end{aligned} \tag{3.1}$$

where $C = q^{-n}/\rho_r(\Lambda)$ is a normalizing value that does not depend on $\bar{\mathbf{a}}$ or $\bar{\mathbf{f}}$. Therefore, by summing over all $\bar{\mathbf{a}}, \bar{\mathbf{f}}$ satisfying $\bar{\mathbf{a}} + \bar{\mathbf{f}} = \bar{\mathbf{v}}$, we obtain that for any $\bar{\mathbf{v}} \in \Lambda/\mathbb{Z}^n$,

$$\Pr[\mathbf{v} = \bar{\mathbf{v}}] \in C\big[1, \tfrac{1+\varepsilon}{1-\varepsilon}\big] \cdot \rho_r(q^{-1}\mathbb{Z}^n + \bar{\mathbf{v}}).$$

Since $r \geq \eta_\varepsilon(q^{-1}\mathbb{Z}^n)$ (by Lemma 2.5), Lemma 2.7 implies that $\Pr[\mathbf{v} = \bar{\mathbf{v}}] \in \big[\frac{1-\varepsilon}{1+\varepsilon}, \frac{1+\varepsilon}{1-\varepsilon}\big]C'$ for a constant C' that is independent of $\bar{\mathbf{v}}$. By Claim 2.1, this shows that $\mathbf{a}'$ is within statistical distance $1-((1-\varepsilon)/(1+\varepsilon))^2 \leq 4\varepsilon$ of the uniform distribution.

It remains to show that the reduction maps $A_{q,\mathbf{s},D_\alpha}$ to $A_{q',\mathbf{G}\mathbf{s},D_\beta}$. Let the input sample from the former distribution be $(\mathbf{a}, b = \langle \mathbf{a}, \mathbf{s}\rangle + e)$, where $e \leftarrow D_\alpha$. As argued above, the output $\mathbf{a}'$ is (nearly) uniform over $\mathbb{T}_{q'}^{n'}$. So condition now on any fixed value $\overline{\mathbf{a}'} \in \mathbb{T}_{q'}^{n'}$ of $\mathbf{a}'$, and let $\bar{\mathbf{v}} = \mathbf{G}^T\overline{\mathbf{a}'} \bmod \mathbb{Z}^n$. We have

$$b' = \langle \mathbf{a}, \mathbf{s}\rangle + e + e' = \langle \overline{\mathbf{a}'}, \mathbf{G}\mathbf{s}\rangle + e + \langle -\mathbf{f}, \mathbf{s}\rangle + e' \bmod 1.$$

By Claim 2.1 and (3.1) (and noting that if $\mathbf{f} = \bar{\mathbf{f}}$ then $\mathbf{a} = \bar{\mathbf{v}} - \bar{\mathbf{f}} \bmod \mathbb{Z}^n$), the distribution of $-\mathbf{f}$ is within statistical distance $1 - (1-\varepsilon)/(1+\varepsilon) \leq 2\varepsilon$ of $D_{q^{-1}\mathbb{Z}^n - \bar{\mathbf{v}}, r}$. By Lemma 2.9 (using $r \geq \sqrt{2}\eta_\varepsilon(q^{-1}\mathbb{Z}^n)$ and $\|\mathbf{s}\| \leq B$), the distribution of $\langle -\mathbf{f}, \mathbf{s}\rangle + e'$ is within statistical distance 6ε from D_t, where $t^2 = r^2(\|\mathbf{s}\|^2 + B^2)$. It therefore follows that $e + \langle -\mathbf{f}, \mathbf{s}\rangle + e'$ is within statistical distance 6ε from $D_{(t^2+\alpha^2)^{1/2}}$, as required. □

4 Hardness of LWE with Binary Secret

The following is the main theorem of this section.

Theorem 4.1. *Let $k, q \geq 1$, and $m \geq n \geq 1$ be integers, and let $\varepsilon \in (0, 1/2)$, $\alpha, \delta > 0$, be such that $n \geq (k+1)\log_2 q + 2\log_2(1/\delta)$, $\alpha \geq \sqrt{\ln(2n(1+1/\varepsilon))/\pi}/q$. There exist three (transformation) reductions from $\mathsf{LWE}_{k,m,q,\alpha}$ to $\mathsf{binLWE}_{n,m,q,\leq\sqrt{10n}\alpha}$, such that for any algorithm for the latter problem with advantage ζ, at least one of the reductions produces an algorithm for the former problem with advantage at least*

$$(\zeta - \delta)/(3m) - 41\varepsilon/2 - \sum_{p|q,\ p \text{ prime}} p^{-k-1}. \tag{4.1}$$

By combining Theorem 4.1 with the reduction in Corollary 3.4 (and noting that $\{0,1\}^n$ is $(\sqrt{n}, 0)$ bounded), we can replace the binLWE problem above with $\mathsf{binLWE}_{n,m,q',\beta}$ for any $q' \geq 1$ and $\xi > 0$ where

$$\beta := \left(10n\alpha^2 + \frac{4n}{\pi q'^2}\ln(2n(1+1/\xi))\right)^{1/2},$$

while decreasing the advantage in (4.1) by $14\xi m$. Recalling that LWE of dimension $k = \sqrt{n}$ and modulus $q = 2^{k/2}$ (assume k is even) is known to be classically as hard as $\sqrt{n}$-dimensional lattice problems (Theorems 2.14 and 2.15), this gives a formal statement of Theorem 1.1. The modulus q' can be taken almost as small as $\sqrt{n}$.

For most purposes the sum over prime factors of q in (4.1) is negligible. For instance, in deriving the formal statement of Theorem 1.1 above, we used a q that is a power of 2, in which case the sum is $2^{-k-1} = 2^{-\sqrt{n}-1}$, which is negligible. If needed, one can improve this by applying the modulus switching reduction (Corollary 3.3) before applying Theorem 4.1 in order to make q prime. (Strictly speaking, one also needs to apply Lemma 2.13 to replace the "unknown noise" variant of LWE given by Corollary 3.3 with the fixed noise variant.) This improves the advantage loss to $q^{-\sqrt{n}-1}$ which is roughly 2^{-n}.

In a high level, the proof of the theorem follows by combining three main steps. The first, given in Section 4.1, reduces LWE to a variant in which the first equation is errorless. The second, given in Section 4.2, reduces the latter to the intermediate problem extLWE, another variant of LWE in which some information on the noise elements is leaked. Finally, in Section 4.3, we reduce extLWE to LWE with $\{0, 1\}$ secret. We note that the first reduction is relatively standard; it is the other two that we consider as the main contribution of this section. We now proceed with more details (see also Figure 1).

Proof. First, since $m \geq n$, Lemma 4.3 provides a transformation reduction from $\mathsf{LWE}_{k,m,q,\alpha}$ to first-is-errorless $\mathsf{LWE}_{k+1,n,q,\alpha}$, while reducing the advantage by at most 2^{-k+1}. Next, Lemma 4.7 with $\mathcal{Z} = \{0, 1\}^n$, which is of quality $\xi = 2$ by Claim 4.6, reduces the latter problem to $\mathsf{extLWE}_{k+1,n,q,\sqrt{5}\alpha,\{0,1\}^n}$ while reducing the advantage by at most $33\varepsilon/2$. Then, Lemma 4.8 reduces the latter problem to $\mathsf{extLWE}^m_{k+1,n,q,\sqrt{5}\alpha,\{0,1\}^n}$, while losing a factor of m in the advantage. Finally, Lemma 4.9 provides three reductions to $\mathsf{binLWE}_{n,m,q,\leq\sqrt{10n}\alpha}$: two from the latter problem, and one from $\mathsf{LWE}_{k+1,m,q,\sqrt{5n}\alpha}$, guaranteeing that the sum of advantages is at least the original advantage minus $4m\varepsilon + \delta$. Together with the trivial reduction from $\mathsf{LWE}_{k,m,q,\alpha}$ to $\mathsf{LWE}_{k+1,m,q,\sqrt{5n}\alpha}$ (which incurs no loss in advantage), this completes the proof. □

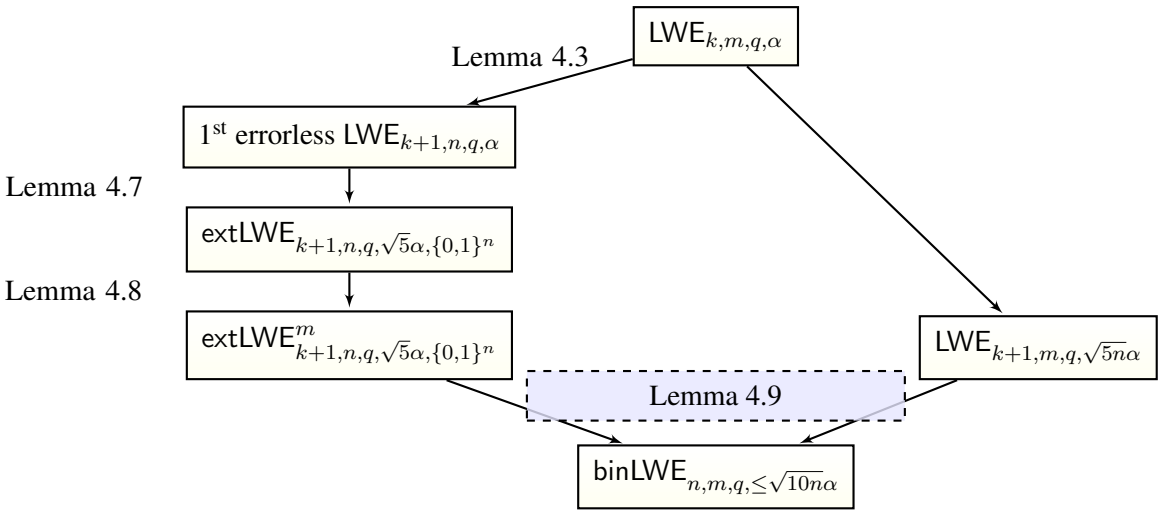

Figure 1: Summary of reductions used in Theorem 4.1

4.1 First-is-errorless LWE

We first define a variant of LWE in which the first equation is given without error, and then show in Lemma 4.3 that it is still hard.

Definition 4.2. *For integers $n, q \geq 1$ and an error distribution ϕ over $\mathbb{R}$, the "first-is-errorless" variant of the* LWE *problem is to distinguish between the following two scenarios. In the first, the first sample is uniform over $\mathbb{T}_q^n \times \mathbb{T}_q$ and the rest are uniform over $\mathbb{T}_q^n \times \mathbb{T}$. In the second, there is an unknown uniformly distributed $\mathbf{s} \in \{0, \ldots, q-1\}^n$, the first sample we get is from $A_{q,\mathbf{s},\{0\}}$ (where $\{0\}$ denotes the distribution that is deterministically zero) and the rest are from $A_{q,\mathbf{s},\phi}$.*

Lemma 4.3. *For any $n \geq 2$, $m, q \geq 1$, and error distribution ϕ, there is an efficient (transformation) reduction from $\mathsf{LWE}_{n-1,m,q,\phi}$ to the first-is-errorless variant of $\mathsf{LWE}_{n,m,q,\phi}$ that reduces the advantage by at most $\sum_p p^{-n}$, with the sum going over all prime factors of q.*

Notice that if q is prime the loss in advantage is at most q^{-n}. Alternatively, for any number q we can bound it by

$$\sum_{k \geq 2} k^{-n} \leq 2^{-n} + \int_2^\infty t^{-n} \mathrm{d}t \leq 2^{-n+2},$$

which might be good enough when n is large.

Proof. The reduction starts by choosing a vector $\mathbf{a}'$ uniformly at random from $\{0, \ldots, q-1\}^n$. Let r be the greatest common divisor of the coordinates of $\mathbf{a}'$. If it is not coprime to q, we abort. The probability that this happens is at most

$$\sum_{p \text{ prime},\ p \mid q} p^{-n}.$$

Assuming we do not abort, we proceed by finding a matrix $\mathbf{U} \in \mathbb{Z}^{n \times n}$ that is invertible modulo q and whose leftmost column is $\mathbf{a}'$. Such a matrix exists, and can be found efficiently. For instance, using the extended GCD algorithm, we find an $n \times n$ unimodular matrix $\mathbf{R}$ such that $\mathbf{R}\mathbf{a}' = (r, 0, \ldots, 0)^T$. Then $\mathbf{R}^{-1} \cdot \mathrm{diag}(r, 1, \ldots, 1)$ is the desired matrix. We also pick a uniform element $s_0 \in \{0, \ldots, q-1\}$. The reduction now proceeds as follows. The first sample it outputs is $(\mathbf{a}'/q, s_0/q)$. The remaining samples are produced by taking a sample $(\mathbf{a}, b)$ from the given oracle, picking a fresh uniformly random $d \in \mathbb{T}_q$, and outputting $(\mathbf{U}(d|\mathbf{a}), b + (s_0 \cdot d))$ with the vertical bar denoting concatenation. It is easy to verify correctness: given uniform samples, the reduction outputs uniform samples (with the first sample's b component uniform over $\mathbb{T}_q$), up to statistical distance 2^{-n+1}; and given samples from $A_{q,\mathbf{s},\phi}$, the reduction outputs one sample from $A_{q,\mathbf{s}',\{0\}}$ and the remaining samples from $A_{q,\mathbf{s}',\phi}$, up to statistical distance 2^{-n+1}, where $\mathbf{s}' = (\mathbf{U}^{-1})^T(s_0|\mathbf{s}) \bmod q$. This proves correctness since $\mathbf{U}$, being invertible modulo q, induces a bijection on $\mathbb{Z}_q^n$, and so $\mathbf{s}'$ is uniform in $\{0, \ldots, q-1\}^n$. □

4.2 Extended LWE

We next define the intermediate problem extLWE. (This definition is of an easier problem than the one considered in previous work [AP12], which makes our hardness result stronger.)

Definition 4.4. *For $n, m, q, t \geq 1$, $\mathcal{Z} \subseteq \mathbb{Z}^m$, and a distribution χ over $\frac{1}{q}\mathbb{Z}^m$, the $\mathsf{extLWE}^t_{n,m,q,\chi,\mathcal{Z}}$ problem is as follows. The algorithm gets to choose $\mathbf{z} \in \mathcal{Z}$ and then receives a tuple*

$$(\mathbf{A}, (\mathbf{b}_i)_{i \in [t]}, (\langle \mathbf{e}_i, \mathbf{z} \rangle)_{i \in [t]}) \in \mathbb{T}_q^{n \times m} \times (\mathbb{T}_q^m)^t \times (\tfrac{1}{q}\mathbb{Z})^t.$$

Its goal is to distinguish between the following two cases. In the first, $\mathbf{A} \in \mathbb{T}_q^{n\times m}$ *is chosen uniformly,* $\mathbf{e}_i \in \frac{1}{q}\mathbb{Z}^m$ *are chosen from* χ*, and* $\mathbf{b}_i = \mathbf{A}^T\mathbf{s}_i + \mathbf{e}_i \bmod 1$ *where* $\mathbf{s}_i \in \{0,\ldots,q-1\}^n$ *are chosen uniformly. The second case is identical, except that the* $\mathbf{b}_i$ *are chosen uniformly in* $\mathbb{T}_q^m$ *independently of everything else.*

When $t = 1$, we omit the superscript t. Also, when χ is $D_{q^{-1}\mathbb{Z}^m,\alpha}$ for some $\alpha > 0$, we replace the subscript χ by α. We note that a discrete version of LWE can be defined as a special case of extLWE by setting $\mathcal{Z} = \{0^m\}$. We next define a measure of quality of sets $\mathcal{Z}$.

Definition 4.5. *For a real* $\xi > 0$ *and a set* $\mathcal{Z} \subseteq \mathbb{Z}^m$ *we say that* $\mathcal{Z}$ *is of* quality ξ *if given any* $\mathbf{z} \in \mathcal{Z}$*, we can efficiently find a unimodular matrix* $\mathbf{U} \in \mathbb{Z}^{m\times m}$ *such that if* $\mathbf{U}' \in \mathbb{Z}^{m\times(m-1)}$ *is the matrix obtained from* $\mathbf{U}$ *by removing its leftmost column then all of the columns of* $\mathbf{U}'$ *are orthogonal to* $\mathbf{z}$ *and its largest singular value is at most* ξ*.*

The idea in this definition is that the columns of $\mathbf{U}'$ form a basis of the lattice of integer points that are orthogonal to $\mathbf{z}$, i.e., the lattice $\{\mathbf{b} \in \mathbb{Z}^m : \langle \mathbf{b}, \mathbf{z}\rangle = 0\}$. The quality measures how "short" we can make this basis.

Claim 4.6. *The set* $\mathcal{Z} = \{0,1\}^m$ *is of quality* 2*.*

Proof. Let $\mathbf{z} \in \mathcal{Z}$ and assume without loss of generality that its first $k \geq 1$ coordinates are 1 and the remaining $m - k$ are 0. Then consider the upper bidiagonal matrix $\mathbf{U}$ whose diagonal is all 1s and whose diagonal above the main diagonal is $(-1,\ldots,-1,0,\ldots,0)$ with -1 appearing $k-1$ times. The matrix is clearly unimodular and all the columns except the first one are orthogonal to $\mathbf{z}$. Moreover, by the triangle inequality, we can bound the operator norm of $\mathbf{U}$ by the sum of that of the diagonal 1 matrix and the off-diagonal matrix, both of which clearly have norm at most 1. □

Lemma 4.7. *Let* $\mathcal{Z} \subseteq \mathbb{Z}^m$ *be of quality* $\xi > 0$*. Then for any* $n, q \geq 1$*,* $\varepsilon \in (0, 1/2)$*, and* $\alpha, r \geq (\ln(2m(1 + 1/\varepsilon))/\pi)^{1/2}/q$*, there is a (transformation) reduction from the first-is-errorless variant of* $\mathsf{LWE}_{n,m,q,\alpha}$ *to* $\mathsf{extLWE}_{n,m,q,(\alpha^2\xi^2+r^2)^{1/2},\mathcal{Z}}$ *that reduces the advantage by at most* $33\varepsilon/2$*.*

Proof. We first describe the reduction. Assume we are asked to provide samples for some $\mathbf{z} \in \mathcal{Z}$. We compute a unimodular $\mathbf{U} \in \mathbb{Z}^{m\times m}$ for $\mathbf{z}$ as in Definition 4.5, and let $\mathbf{U}' \in \mathbb{Z}^{m\times(m-1)}$ be the matrix formed by removing the first column of $\mathbf{U}$. We then take m samples from the given distribution, resulting in $(\mathbf{A}, \mathbf{b}) \in \mathbb{T}_q^{n\times m} \times (\mathbb{T}_q \times \mathbb{T}^{m-1})$. We also sample a vector $\mathbf{f}$ from the m-dimensional continuous Gaussian distribution $D_{\alpha(\xi^2\mathbf{I}-\mathbf{U}'\mathbf{U}'^T)^{1/2}}$, which is well defined since $\xi^2\mathbf{I} - \mathbf{U}'\mathbf{U}'^T$ is a positive semidefinite matrix by our assumption on $\mathbf{U}$. The output of the reduction is the tuple

$$(\mathbf{A}' = \mathbf{A}\mathbf{U}^T, \mathbf{b}' + \mathbf{c}, \langle \mathbf{z}, \mathbf{f} + \mathbf{c}\rangle) \in \mathbb{T}_q^{n\times m} \times \mathbb{T}_q^m \times \tfrac{1}{q}\mathbb{Z}, \tag{4.2}$$

where $\mathbf{b}' = \mathbf{U}\mathbf{b} + \mathbf{f}$, and $\mathbf{c}$ is chosen from the discrete Gaussian distribution $D_{q^{-1}\mathbb{Z}^m-\mathbf{b}',r}$ (using Lemma 2.3). (As before, notice that the coset $q^{-1}\mathbb{Z}^m - \mathbf{b}'$ is well defined because $\mathbf{b}'$ is a coset of $\mathbb{Z}^m \subseteq q^{-1}\mathbb{Z}^m$.)

We now prove the correctness of the reduction. Consider first the case that we get valid LWE equations, i.e., $\mathbf{A}$ is uniform in $\mathbb{T}_q^{n\times m}$ and $\mathbf{b} = \mathbf{A}^T\mathbf{s} + \mathbf{e} \in \mathbb{T}^m$ where $\mathbf{s} \in \{0,\ldots,q-1\}^n$ is uniformly chosen, the first coordinate of $\mathbf{e} \in \mathbb{R}^m$ is 0, and the remaining $m-1$ coordinates are chosen from D_α. Since $\mathbf{U}$ is unimodular, $\mathbf{A}' = \mathbf{A}\mathbf{U}^T$ is uniformly distributed in $\mathbb{T}_q^{n\times m}$ as required. From now on we condition on an arbitrary $\mathbf{A}'$ and analyze the distribution of the remaining two components of (4.2). Next,

$$\mathbf{b}' = \mathbf{U}\mathbf{b} + \mathbf{f} = \mathbf{A}'^T\mathbf{s} + \mathbf{U}\mathbf{e} + \mathbf{f}.$$

Since $\mathbf{Ue}$ is distributed as a continuous Gaussian $D_{\alpha\mathbf{U}'}$, the vector $\mathbf{Ue}+\mathbf{f}$ is a distributed as a *spherical* continuous Gaussian $D_{\alpha\xi}$. Moreover, since $\mathbf{A}'^T\mathbf{s} \in \mathbb{T}_q^m$, the coset $q^{-1}\mathbb{Z}^m - \mathbf{b}'$ is identical to $q^{-1}\mathbb{Z}^m - (\mathbf{Ue}+\mathbf{f})$, so we can see $\mathbf{c}$ as being chosen from $D_{q^{-1}\mathbb{Z}^m-(\mathbf{Ue}+\mathbf{f}),r}$. Therefore, by Lemma 2.10 and using that $r \geq \eta_\varepsilon(q^{-1}\mathbb{Z}^m)$ by Lemma 2.5, the distribution of $\mathbf{Ue}+\mathbf{f}+\mathbf{c}$ is within statistical distance 8ε of $D_{q^{-1}\mathbb{Z}^m,(\alpha^2\xi^2+r^2)^{1/2}}$. This shows that the second component (4.2) is also distributed correctly. Finally, for the third component, by our assumption on $\mathbf{U}$ and the fact that the first coordinate of $\mathbf{e}$ is zero,

$$\langle \mathbf{z}, \mathbf{f}+\mathbf{c}\rangle = \langle \mathbf{z}, \mathbf{Ue}+\mathbf{f}+\mathbf{c}\rangle,$$

and so the third component gives the inner product of the noise with $\mathbf{z}$, as desired.

We now consider the case where the input is uniform, i.e., that $\mathbf{A}$ is uniform in $\mathbb{T}_q^{n\times m}$ and $\mathbf{b}$ is independent and uniform in $\mathbb{T}_q \times \mathbb{T}^{m-1}$. We first observe that by Lemma 2.6, since $\alpha \geq \eta_{\varepsilon/m}(q^{-1}\mathbb{Z})$ (by Lemma 2.5), the distribution of $(\mathbf{A},\mathbf{b})$ is within statistical distance $\varepsilon/2$ of the distribution of $(\mathbf{A},\mathbf{e}'+\mathbf{e})$ where $\mathbf{e}'$ is chosen uniformly in $\mathbb{T}_q^m$, the first coordinate of $\mathbf{e}$ is zero, and its remaining $m-1$ coordinates are chosen independently from D_α. So from now on assume our input is $(\mathbf{A},\mathbf{e}'+\mathbf{e})$. The first component of (4.2) is uniform in $\mathbb{T}_q^{n\times m}$ as before, and moreover, it is clearly independent of the other two. Moreover, since $\mathbf{b}' = \mathbf{Ue}'+\mathbf{Ue}+\mathbf{f}$ and $\mathbf{Ue}' \in \mathbb{T}_q^m$, the coset $q^{-1}\mathbb{Z}^m - \mathbf{b}'$ is identical to $q^{-1}\mathbb{Z}^m - (\mathbf{Ue}+\mathbf{f})$, and so $\mathbf{c}$ is distributed identically to the case of a valid LWE equation, and in particular is independent of $\mathbf{e}'$. This establishes that the third component of (4.2) is correctly distributed; moreover, since $\mathbf{e}'$ is independent of the first and third components, and $\mathbf{Ue}'$ is uniform in $\mathbb{T}_q^m$ (since $\mathbf{U}$ is unimodular), we get that the second component is uniform and independent of the other two, as desired. □

We end this section by stating the standard reduction to the multi-secret ($t \geq 1$) case of extended LWE.

Lemma 4.8. *Let $n, m, q, \chi, \mathcal{Z}$ be as in Definition 4.4 with χ efficiently sampleable, and let $t \geq 1$ be an integer. Then there is an efficient (transformation) reduction from $\mathsf{extLWE}_{n,m,q,\chi,\mathcal{Z}}$ to $\mathsf{extLWE}^t_{n,m,q,\chi,\mathcal{Z}}$ that shrinks the advantage by a factor of t.*

The proof is by a standard hybrid argument. We bring it here for the sake of completeness. We note that the distribution of the secret vector $\mathbf{s}$ needs to be sampleable but otherwise it plays no role in the proof. The lemma therefore naturally extends to any (sampleable) distribution of $\mathbf{s}$.

Proof. Let $\mathcal{A}$ be an algorithm for $\mathsf{extLWE}^t_{n,m,q,\chi,\mathcal{Z}}$, let $\mathbf{z}$ be the vector output by $\mathcal{A}$ in the first step (note that this is a random variable) and let H_i denote the distribution

$$\left(\mathbf{A}, \{\mathbf{b}_1,\ldots,\mathbf{b}_i,\mathbf{u}_{i+1},\ldots,\mathbf{u}_t\}, \mathbf{z}, \{\langle \mathbf{z},\mathbf{e}_i\rangle\}_{i\in[t]}\right) ,$$

where $\mathbf{u}_{i+1},\ldots,\mathbf{u}_t$ are sampled independently and uniformly in $\mathbb{T}_q^m$. Then by definition $\mathrm{Adv}[\mathcal{A}] = |\Pr[\mathcal{A}(H_0)] - \Pr[\mathcal{A}(H_t)]|$.

We now describe an algorithm $\mathcal{B}$ for $\mathsf{extLWE}_{n,m,q,\chi,\mathcal{Z}}$: First, $\mathcal{B}$ runs $\mathcal{A}$ to obtain $\mathbf{z}$ and sends it to the challenger as its own $\mathbf{z}$. Then, given an input $(\mathbf{A},\mathbf{d},\mathbf{z},y)$ for $\mathsf{extLWE}_{n,m,q,\chi,\mathcal{Z}}$, the distinguisher $\mathcal{B}$ samples $i^* \leftarrow [t]$, and in addition $\mathbf{s}_1,\ldots,\mathbf{s}_{i^*-1} \leftarrow \mathbb{Z}_q^n$, $\mathbf{e}_1,\ldots,\mathbf{e}_{i^*-1},\mathbf{e}_{i^*+1},\ldots,\mathbf{e}_t \leftarrow \chi^m$, $\mathbf{u}_{i^*+1},\ldots,\mathbf{u}_t \leftarrow \mathbb{T}_q^m$. It sets $\mathbf{b}_i = \mathbf{A}^T\cdot\mathbf{s}_i + \mathbf{e}_i \pmod 1$, and sends the following to $\mathcal{A}$:

$$(\mathbf{A}, \{\mathbf{b}_1,\ldots,\mathbf{b}_{i^*-1},\mathbf{d},\mathbf{u}_{i^*+1},\ldots,\mathbf{u}_t\}, \mathbf{z}, \{\langle \mathbf{z},\mathbf{e}_1\rangle,\ldots,\langle \mathbf{z},\mathbf{e}_{i^*-1}\rangle, y, \langle \mathbf{z},\mathbf{e}_{i^*+1}\rangle,\ldots,\langle \mathbf{z},\mathbf{e}_t\rangle\}) .$$

Finally, $\mathcal{B}$ outputs the same output as $\mathcal{A}$ did.

Note that when the input to $\mathcal{B}$ is distributed as $P_0 = (\mathbf{A}, \mathbf{b}, \mathbf{z}, \mathbf{z}^T \cdot \mathbf{e})$ with $\mathbf{b} = \mathbf{A}^T \cdot \mathbf{s} + \mathbf{e} \pmod 1$, then $\mathcal{B}$ feeds $\mathcal{A}$ with exactly the distribution H_{i^*}. On the other hand, if the input to $\mathcal{B}$ is $P_1 = (\mathbf{A}, \mathbf{u}, \mathbf{z}, \mathbf{z}^T \cdot \mathbf{e})$ with $\mathbf{u} \leftarrow \mathbb{T}_q^m$, then $\mathcal{B}$ feeds $\mathcal{A}$ with H_{i^*-1}.

Since i^* is uniform in $[t]$, we get that

$$\begin{aligned} t \operatorname{Adv}[\mathcal{B}] &= t \left|\Pr[\mathcal{B}(P_0)] - \Pr[\mathcal{B}(P_1)]\right| \\ &= \left|\sum_{i^* \in [t]} \Pr[\mathcal{A}(H_{i^*})] - \sum_{i^* \in [t]} \Pr[\mathcal{A}(H_{i^*-1})]\right| \\ &= \left|\Pr[\mathcal{A}(H_t)] - \Pr[\mathcal{A}(H_0)]\right| \\ &= \operatorname{Adv}[\mathcal{A}] \,, \end{aligned}$$

and the result follows. □

4.3 Reducing to binary secret

Lemma 4.9. *Let $k, n, m, q \in \mathbb{N}$, $\varepsilon \in (0, 1/2)$, and $\delta, \alpha, \beta, \gamma > 0$ be such that $n \geq k \log_2 q + 2\log_2(1/\delta)$, $\beta \geq \sqrt{2\ln(2n(1+1/\varepsilon))/\pi}/q$, $\alpha = \sqrt{2n}\beta$, $\gamma = \sqrt{n}\beta$. Then there exist three efficient (transformation) reductions to $\mathsf{binLWE}_{n,m,q,\leq\alpha}$ from $\mathsf{extLWE}^m_{k,n,q,\beta,\{0,1\}^n}$, $\mathsf{LWE}_{k,m,q,\gamma}$, and $\mathsf{extLWE}^m_{k,n,q,\beta,\{0^n\}}$, such that if $\mathcal{B}_1$, $\mathcal{B}_2$, and $\mathcal{B}_3$ are the algorithms obtained by applying these reductions (respectively) to an algorithm $\mathcal{A}$, then*

$$\operatorname{Adv}[\mathcal{A}] \leq \operatorname{Adv}[\mathcal{B}_1] + \operatorname{Adv}[\mathcal{B}_2] + \operatorname{Adv}[\mathcal{B}_3] + 4m\varepsilon + \delta \,.$$

Pointing out the trivial (transformation) reduction from $\mathsf{extLWE}^m_{k,n,q,\beta,\{0,1\}^n}$ to $\mathsf{extLWE}^m_{k,n,q,\beta,\{0^n\}}$, the lemma implies the hardness of $\mathsf{binLWE}_{n,m,q,\leq\alpha}$ based on the hardness of $\mathsf{extLWE}^m_{k,n,q,\beta,\{0,1\}^n}$ and $\mathsf{LWE}_{k,m,q,\gamma}$.

We note that our proof is actually more general, and holds for any binary distribution of min-entropy at least $k \log_2 q + 2\log_2(1/\delta)$, and not just a uniform binary secret as in the definition of binLWE.

Proof. The proof follows by a sequence of hybrids. Let $k, n, m, q, \varepsilon, \alpha, \beta, \gamma$ be as in the lemma statement. We consider $\mathbf{z} \leftarrow \{0,1\}^n$ and $\mathbf{e} \leftarrow D^m_{\alpha'}$ for $\alpha' = \sqrt{\beta^2\|\mathbf{z}\|^2 + \gamma^2} \leq \sqrt{2n}\beta = \alpha$. In addition, we let $\mathbf{A} \leftarrow \mathbb{T}_q^{n \times m}$, $\mathbf{u} \leftarrow \mathbb{T}^m$, and define $\mathbf{b} := \mathbf{A}^T \cdot \mathbf{z} + \mathbf{e} \pmod 1$. We consider an algorithm $\mathcal{A}$ that distinguishes between $(\mathbf{A}, \mathbf{b})$ and $(\mathbf{A}, \mathbf{u})$.

We let H_0 denote the distribution $(\mathbf{A}, \mathbf{b})$ and H_1 the distribution

$$H_1 = (\mathbf{A}, \mathbf{A}^T\mathbf{z} - \mathbf{N}^T\mathbf{z} + \hat{\mathbf{e}} \bmod 1),$$

where $\mathbf{N} \leftarrow D^{n \times m}_{q^{-1}\mathbb{Z},\beta}$ and $\hat{\mathbf{e}} \leftarrow D^m_\gamma$. Using $\|\mathbf{z}\| \leq \sqrt{n}$ and that $\beta \geq \sqrt{2}\eta_\varepsilon(\mathbb{Z}^n)/q$ (by Lemma 2.5), it follows by Lemma 2.9 that the statistical distance between $-\mathbf{N}^T\mathbf{z} + \hat{\mathbf{e}}$ and $D^m_{\alpha'}$ is at most $4m\varepsilon$. It thus follows that

$$\left|\Pr[\mathcal{A}(H_0)] - \Pr[\mathcal{A}(H_1)]\right| \leq 4m\varepsilon \,. \tag{4.3}$$

We define a distribution H_2 as follows. Let $\mathbf{B} \leftarrow \mathbb{T}_q^{k \times m}$ and $\mathbf{C} \leftarrow \mathbb{T}_q^{k \times n}$. Let $\hat{\mathbf{A}} := q\mathbf{C}^T \cdot \mathbf{B} + \mathbf{N} \pmod 1$. Finally,

$$H_2 = (\hat{\mathbf{A}}, \hat{\mathbf{A}}^T \cdot \mathbf{z} - \mathbf{N}^T\mathbf{z} + \hat{\mathbf{e}}) = (\hat{\mathbf{A}}, q\mathbf{B}^T \cdot \mathbf{C} \cdot \mathbf{z} + \hat{\mathbf{e}}) \,.$$

We now argue that there exists an adversary $\mathcal{B}_1$ for problem $\mathsf{extLWE}^m_{k,n,q,\beta,\{0,1\}^n}$, such that

$$\operatorname{Adv}[\mathcal{B}_1] = \left|\Pr[\mathcal{A}(H_1)] - \Pr[\mathcal{A}(H_2)]\right| \,. \tag{4.4}$$

This is because H_1, H_2 can be viewed as applying the same efficient transformation on the distributions $(\mathbf{C}, \mathbf{A}, \mathbf{N}^T\mathbf{z})$ and $(\mathbf{C}, \hat{\mathbf{A}}, \mathbf{N}^T\mathbf{z})$ respectively. Since distinguishing the latter distributions is exactly the $\mathsf{extLWE}^m_{k,n,q,\beta,\{0,1\}^n}$ problem (where the columns of $q \cdot \mathbf{B}$ are interpreted as the m secret vectors), the distinguisher $\mathcal{B}_1$ follows by first applying the aforementioned transformation and then applying $\mathcal{A}$.

For the next hybrid, we define $H_3 = (\hat{\mathbf{A}}, \mathbf{B}^T \cdot \mathbf{s} + \hat{\mathbf{e}})$, for $\mathbf{s} \leftarrow \mathbb{Z}_q^k$. It follows that

$$|\Pr[\mathcal{A}(H_2)] - \Pr[\mathcal{A}(H_3)]| \leq \delta \tag{4.5}$$

by the leftover hash lemma (see Lemma 2.2), since H_2, H_3 can be derived from $(\mathbf{C}, q\mathbf{C} \cdot \mathbf{z})$ and $(\mathbf{C}, \mathbf{s})$ respectively, whose statistical distance is at most δ.

Our next hybrid makes the second component uniform: $H_4 = (\hat{\mathbf{A}}, \mathbf{u})$. There exists an algorithm $\mathcal{B}_2$ for $\mathsf{LWE}_{k,m,q,\gamma}$ such that

$$\mathrm{Adv}[\mathcal{B}_2] = |\Pr[\mathcal{A}(H_3)] - \Pr[\mathcal{A}(H_4)]| \,, \tag{4.6}$$

since H_3, H_4 can be computed efficiently from $(\mathbf{B}, \mathbf{B}^T\mathbf{s} + \hat{\mathbf{e}}), (\mathbf{B}, \mathbf{u})$.

Lastly, we change $\hat{\mathbf{A}}$ back to uniform: $H_5 = (\mathbf{A}, \mathbf{u})$. There exists an algorithm $\mathcal{B}_3$ for $\mathsf{extLWE}^m_{k,n,q,\beta,\{0^n\}}$ such that

$$\mathrm{Adv}[\mathcal{B}_3] = |\Pr[\mathcal{A}(H_4)] - \Pr[\mathcal{A}(H_5)]| \,. \tag{4.7}$$

Eq. (4.7) is derived very similarly to Eq. (4.4): We notice that H_4, H_5 can be viewed as applying the same efficient transformation on the distributions $(\mathbf{C}, \hat{\mathbf{A}})$ and $(\mathbf{C}, \mathbf{A})$ respectively. Since distinguishing the latter distributions is exactly the $\mathsf{extLWE}^m_{k,n,q,\beta,\{0^n\}}$ problem (where the columns of $q \cdot \mathbf{B}$ are interpreted as the m secret vectors), the distinguisher $\mathcal{B}_3$ follows by first applying the aforementioned transformation and then applying $\mathcal{A}$.

Putting together Eq. (4.3), (4.4), (4.5), (4.6), (4.7), the lemma follows. □

5 Exact Gaussian Sampler

In this section we prove Lemma 2.3. As in [GPV08], the proof consists of two parts. In the first we consider the one-dimensional case, and in the second we use it recursively to sample from arbitrary lattices. Our one-dimensional sampler is based on rejection sampling, just like the one in [GPV08]. Unlike [GPV08], we use the continuous normal distribution as the source distribution which allows us to avoid truncation, and as a result obtain an exact sample. Our second part uses the same recursive routine as in [GPV08], but adds a rejection sampling step to it in order to take care of the deviation of its output from the desired distribution.

Before proceeding to the proof, we remark that another approach towards improving the sampler of [GPV08] would consist in adapting the Markov chain Monte Carlo sampler of Sinclair and Jerrum [SJ89] (see also [HS10]). We expect that this would allow to improve the bound on r in Lemma 2.3 to $\|\widetilde{\mathbf{B}}\|$ at the expense of a larger (but still polynomial) running time.[1] Because this extra effort seems to offer small gain, we do not pursue this direction further here.

5.1 The one-dimensional case

Here we show how to sample from the discrete Gaussian distribution on arbitrary cosets of one-dimensional lattices. We use a standard rejection sampling procedure (see, e.g. [Dev86, Page 117] for a very similar procedure).

[1] We thank Ronnie Barequet for discussions that led to this idea.

By scaling, we can restrict without loss of generality to the lattice $\mathbb{Z}$, i.e., we consider the task of sampling from $D_{\mathbb{Z}+c,r}$ for a given coset representative $c \in [0,1)$ and parameter $r > 0$. The sampling procedure is as follows. Let $Z_0 = \int_c^\infty \rho_r(x)\mathrm{d}x$, and $Z_1 = \int_{-\infty}^{c-1} \rho_r(x)\mathrm{d}x$. These two numbers can be computed efficiently by expressing them in terms of the error function. Let $Z = Z_0 + Z_1 + \rho_r(c) + \rho_r(c-1)$. The algorithm repeats the following until it outputs an answer:

- With probability $\rho_r(c)/Z$ it outputs c;
- With probability $\rho_r(c-1)/Z$ it outputs $c-1$;
- With probability Z_0/Z it chooses x from the restriction of the continuous normal distribution D_r to the interval $[c,\infty)$. Let y be the smallest element in $\mathbb{Z}+c$ that is larger than x. With probability $\rho_r(y)/\rho_r(x)$ output y, and otherwise repeat;
- With probability Z_1/Z it chooses x from the restriction of the continuous normal distribution D_r to the interval $(-\infty, c-1]$. Let y be the largest element in $\mathbb{Z}+c$ that is smaller than x. With probability $\rho_r(y)/\rho_r(x)$ output y, and otherwise repeat.

Consider now one iteration of the procedure. The probability of outputting c is $\rho_r(c)/Z$, that of outputting $c-1$ is $\rho_r(c-1)/Z$, that of outputting $c+k$ for some $k \geq 1$ is

$$\frac{Z_0}{Z} \cdot \frac{1}{Z_0} \int_{c+k-1}^{c+k} \rho_r(x) \cdot \frac{\rho_r(c+k)}{\rho_r(x)} \mathrm{d}x = \frac{\rho_r(c+k)}{Z},$$

and similarly, that of outputting $c-1-k$ for some $k \geq 1$ is $\rho_r(c-1-k)/Z$. From this it follows immediately that conditioned on outputting something, the output distribution has support on $\mathbb{Z}+c$ and probability mass function proportional to ρ_r, and is therefore the desired discrete Gaussian distribution $D_{\mathbb{Z}+c,r}$. Moreover, the probability of outputting something is

$$\frac{\rho_r(\mathbb{Z}+c)}{Z} = \frac{\rho_r(\mathbb{Z}+c)}{Z_0 + Z_1 + \rho_r(c) + \rho_r(c-1)} \geq \frac{\rho_r(\mathbb{Z}+c)}{\rho_r(\mathbb{Z}+c) + \rho_r(c) + \rho_r(c-1)} \geq \frac{1}{2}.$$

Therefore at each iteration the procedure has probability of at least $1/2$ to terminate. As a result, the probability that the number of iterations is greater than t is at most 2^{-t}, and in particular, the expected number of iterations is at most 2.

5.2 The general case

For completeness, we start by recalling the SampleD procedure described in [GPV08]. This is a recursive procedure that gets as input a basis $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of an n-dimensional lattice $\Lambda = \mathcal{L}(\mathbf{B})$, a parameter $r > 0$, and a vector $\mathbf{c} \in \mathbb{R}^n$, and outputs a vector in $\Lambda + \mathbf{c}$ whose distribution is close to that of $D_{\Lambda+\mathbf{c},r}$. Let $\widetilde{\mathbf{b}_1}, \ldots, \widetilde{\mathbf{b}_n}$ be the Gram-Schmidt orthogonalization of $\mathbf{b}_1, \ldots, \mathbf{b}_n$, and let $\overline{\mathbf{b}_1}, \ldots, \overline{\mathbf{b}_n}$ be the normalized Gram-Schmidt vectors, i.e., $\overline{\mathbf{b}_i} = \widetilde{\mathbf{b}_i}/\|\widetilde{\mathbf{b}_i}\|$. The procedure is the following.

1. Let $\mathbf{c}_n \leftarrow \mathbf{c}$. For $i \leftarrow n, \ldots, 1$, do:
   (a) Choose v_i from $D_{\|\widetilde{\mathbf{b}_i}\|\mathbb{Z}+\langle \mathbf{c}_i, \overline{\mathbf{b}_i}\rangle, r}$ using the exact one-dimensional sampler.
   (b) Let $\mathbf{c}_{i-1} \leftarrow \mathbf{c}_i + (v_i - \langle \mathbf{c}_i, \overline{\mathbf{b}_i}\rangle) \cdot \mathbf{b}_i/\|\widetilde{\mathbf{b}_i}\| - v_i\overline{\mathbf{b}_i}$.

2. Output $\mathbf{v} := \sum_{i=1}^{n} v_i \overline{\mathbf{b}_i}$.

It is easy to verify that the procedure always outputs vectors in the coset $\Lambda + \mathbf{c}$. Moreover, the probability of outputting any $\mathbf{v} \in \Lambda + \mathbf{c}$ is

$$\prod_{i=1}^{n} \frac{\rho_r(v_i)}{\rho_r(\|\widetilde{\mathbf{b}_i}\|\mathbb{Z} + \langle \mathbf{c}_i, \overline{\mathbf{b}_i} \rangle)} = \frac{\rho_r(\mathbf{v})}{\prod_{i=1}^{n} \rho_r(\|\widetilde{\mathbf{b}_i}\|\mathbb{Z} + \langle \mathbf{c}_i, \overline{\mathbf{b}_i} \rangle)},$$

where $\mathbf{c}_i$ are the values computed in the procedure when it outputs $\mathbf{v}$. Notice that by Lemma 2.5 and our assumption on r, we have that $r \geq \eta_{1/(n+1)}(\|\widetilde{\mathbf{b}_i}\|\mathbb{Z})$ for all i. Therefore, by Lemma 2.7, we have that for all $c \in \mathbb{R}$,

$$\rho_r(\|\widetilde{\mathbf{b}_i}\|\mathbb{Z} + c) \in \left[1 - \frac{2}{n+2}, 1\right] \rho_r(\|\widetilde{\mathbf{b}_i}\|\mathbb{Z}).$$

In order to get an exact sample, we combine the above procedure with rejection sampling. Namely, we apply SampleD to obtain some vector $\mathbf{v}$. We then output $\mathbf{v}$ with probability

$$\frac{\prod_{i=1}^{n} \rho_r(\|\widetilde{\mathbf{b}_i}\|\mathbb{Z} + \langle \mathbf{c}_i, \overline{\mathbf{b}_i} \rangle)}{\prod_{i=1}^{n} \rho_r(\|\widetilde{\mathbf{b}_i}\|\mathbb{Z})} \in \left(\left(1 - \frac{2}{n+2}\right)^n, 1\right] \subseteq (e^{-2}, 1], \tag{5.1}$$

and otherwise repeat. This probability can be efficiently computed, as we will show below. As a result, in any given iteration the probability of outputting the vector $\mathbf{v} \in \Lambda + \mathbf{c}$ is

$$\frac{\rho_r(\mathbf{v})}{\prod_{i=1}^{n} \rho_r(\|\widetilde{\mathbf{b}_i}\|\mathbb{Z})}.$$

Since the denominator is independent of $\mathbf{v}$, we obtain that in any given iteration, conditioned on outputting something, the output is distributed according to the desired distribution $D_{\Lambda+\mathbf{c},r}$, and therefore this is also the overall output distribution of our sampler. Moreover, by (5.1), the probability of outputting something in any given iteration is at least e^{-2}, and therefore, the probability that the number of iterations is greater than t is at most $(1 - e^{-2})^t$, and in particular, the expected number of iterations is at most e^2.

It remains to show how to efficiently compute the probability in (5.1). By scaling, it suffices to show how to compute

$$\rho_r(\mathbb{Z} + c) = \sum_{k \in \mathbb{Z}} \exp(-\pi(k + c)^2 / r^2)$$

for any $r > 0$ and $c \in [0, 1)$. If $r < 1$, the sum decays very fast, and we can achieve any desired t bits of accuracy in time $\mathrm{poly}(t)$, which agrees with our notion of efficiently computing a real number (following, e.g., the treatment in [Lov86, Section 1.4]). For $r \geq 1$, we use the Poisson summation formula (see, e.g., [MR04, Lemma 2.8]) to write

$$\rho_r(\mathbb{Z} + c) = r \cdot \sum_{k \in \mathbb{Z}} \exp(-\pi k^2 r^2 + 2\pi i c k) = r \cdot \sum_{k \in \mathbb{Z}} \exp(-\pi k^2 r^2) \cos(2\pi c k),$$

which again decays fast enough so we can compute it to within any desired t bits of accuracy in time $\mathrm{poly}(t)$.

Acknowledgments: We thank Elette Boyle, Adam Klivans, Vadim Lyubashevsky, Sasha Sherstov, Vinod Vaikuntanathan and Gilles Villard for useful discussions.

References

[ABB10a] S. Agrawal, D. Boneh, and X. Boyen. Efficient lattice (H)IBE in the standard model. In *EUROCRYPT*, pages 553–572. 2010.

[ABB10b] S. Agrawal, D. Boneh, and X. Boyen. Lattice basis delegation in fixed dimension and shorter-ciphertext hierarchical IBE. In *CRYPTO*, pages 98–115. 2010.

[ACPS09] B. Applebaum, D. Cash, C. Peikert, and A. Sahai. Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In *CRYPTO*, pages 595–618. 2009.

[AD87] D. Aldous and P. Diaconis. Strong uniform times and finite random walks. *Adv. in Appl. Math.*, 8(1):69–97, 1987. ISSN 0196-8858. doi:10.1016/0196-8858(87)90006-6.

[AD97] M. Ajtai and C. Dwork. A public-key cryptosystem with worst-case/average-case equivalence. In *STOC*, pages 284–293. 1997.

[AGV09] A. Akavia, S. Goldwasser, and V. Vaikuntanathan. Simultaneous hardcore bits and cryptography against memory attacks. In *TCC*, pages 474–495. 2009.

[Ajt96] M. Ajtai. Generating hard instances of lattice problems. In *Complexity of computations and proofs*, volume 13 of *Quad. Mat.*, pages 1–32. Dept. Math., Seconda Univ. Napoli, Caserta, 2004. Preliminary version in STOC 1996.

[AP12] J. Alperin-Sheriff and C. Peikert. Circular and KDM security for identity-based encryption. In *Public Key Cryptography*, pages 334–352. 2012.

[Ban93] W. Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(4):625–635, 1993.

[BGV12] Z. Brakerski, C. Gentry, and V. Vaikuntanathan. (Leveled) fully homomorphic encryption without bootstrapping. In *ITCS*, pages 309–325. 2012.

[Boy10] X. Boyen. Lattice mixing and vanishing trapdoors: A framework for fully secure short signatures and more. In *Public Key Cryptography*, pages 499–517. 2010.

[Bra12] Z. Brakerski. Fully homomorphic encryption without modulus switching from classical GapSVP. In *CRYPTO*, pages 868–886. 2012.

[BV96] D. Boneh and R. Venkatesan. Hardness of computing the most significant bits of secret keys in Diffie-Hellman and related schemes. In N. Koblitz, editor, *CRYPTO*, volume 1109 of *Lecture Notes in Computer Science*, pages 129–142. Springer, 1996. ISBN 3-540-61512-1.

[BV11] Z. Brakerski and V. Vaikuntanathan. Efficient fully homomorphic encryption from (standard) LWE. In *FOCS*, pages 97–106. 2011.

[CHKP10] D. Cash, D. Hofheinz, E. Kiltz, and C. Peikert. Bonsai trees, or how to delegate a lattice basis. In *EUROCRYPT*, pages 523–552. 2010.

[Dev86] L. Devroye. *Nonuniform random variate generation*. Springer-Verlag, New York, 1986. Available at `http://luc.devroye.org/rnbookindex.html`.

[Gen09] C. Gentry. Fully homomorphic encryption using ideal lattices. In *STOC*, pages 169–178. 2009.

[GGH96] O. Goldreich, S. Goldwasser, and S. Halevi. Collision-free hashing from lattice problems. *Electronic Colloquium on Computational Complexity (ECCC)*, 3(42), 1996.

[GHPS12] C. Gentry, S. Halevi, C. Peikert, and N. P. Smart. Ring switching in BGV-style homomorphic encryption. In *SCN*, pages 19–37. 2012.

[GKPV10] S. Goldwasser, Y. T. Kalai, C. Peikert, and V. Vaikuntanathan. Robustness of the learning with errors assumption. In *ICS*, pages 230–240. 2010.

[GPV08] C. Gentry, C. Peikert, and V. Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *STOC*, pages 197–206. 2008.

[HILL99] J. Håstad, R. Impagliazzo, L. A. Levin, and M. Luby. A pseudorandom generator from any one-way function. *SIAM J. Comput.*, 28(4):1364–1396, 1999.

[HS10] T. P. Hayes and A. Sinclair. Liftings of tree-structured Markov chains. In *APPROX-RANDOM*, pages 602–616. 2010.

[Kho10] S. Khot. Inapproximability results for computational problems on lattices. In P. Nguyen and B. Vallée, editors, *The LLL Algorithm: Survey and Applications*. Springer-Verlag, New York, 2010.

[KS06] A. R. Klivans and A. A. Sherstov. Cryptographic hardness for learning intersections of halfspaces. In *FOCS*, pages 553–562. 2006.

[KTX08] A. Kawachi, K. Tanaka, and K. Xagawa. Concurrently secure identification schemes based on the worst-case hardness of lattice problems. In *ASIACRYPT*, pages 372–389. 2008.

[Kup05] G. Kuperberg. A subexponential-time quantum algorithm for the dihedral hidden subgroup problem. *SIAM J. Comput.*, 35(1):170–188, 2005.

[LM06] V. Lyubashevsky and D. Micciancio. Generalized compact knapsacks are collision resistant. In *ICALP (2)*, pages 144–155. 2006.

[LM09] V. Lyubashevsky and D. Micciancio. On bounded distance decoding, unique shortest vectors, and the minimum distance problem. In *CRYPTO*, pages 577–594. 2009.

[Lov86] L. Lovász. *An algorithmic theory of numbers, graphs and convexity*, volume 50 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1986.

[LP11] R. Lindner and C. Peikert. Better key sizes (and attacks) for LWE-based encryption. In *CT-RSA*, pages 319–339. 2011.

[LPR10] V. Lyubashevsky, C. Peikert, and O. Regev. On ideal lattices and learning with errors over rings. In *EUROCRYPT*, pages 1–23. 2010.

[Lyu08] V. Lyubashevsky. Lattice-based identification schemes secure under active attacks. In *Public Key Cryptography*, pages 162–179. 2008.

[Lyu12] V. Lyubashevsky. Lattice signatures without trapdoors. In *EUROCRYPT*, pages 738–755. 2012.

[MM11] D. Micciancio and P. Mol. Pseudorandom knapsacks and the sample complexity of LWE search-to-decision reductions. In *CRYPTO*, pages 465–484. 2011.

[MP12] D. Micciancio and C. Peikert. Trapdoors for lattices: Simpler, tighter, faster, smaller. In *EUROCRYPT*, pages 700–718. 2012.

[MR04] D. Micciancio and O. Regev. Worst-case to average-case reductions based on Gaussian measures. *SIAM J. Comput.*, 37(1):267–302, 2007. Preliminary version in FOCS 2004.

[MV03] D. Micciancio and S. P. Vadhan. Statistical zero-knowledge proofs with efficient provers: Lattice problems and more. In *CRYPTO*, pages 282–298. 2003.

[OPW11] A. O'Neill, C. Peikert, and B. Waters. Bi-deniable public-key encryption. In *CRYPTO*, pages 525–542. 2011.

[Pei09] C. Peikert. Public-key cryptosystems from the worst-case shortest vector problem. In *STOC*, pages 333–342. 2009.

[Pei10] C. Peikert. An efficient and parallel Gaussian sampler for lattices. In *CRYPTO*, pages 80–97. 2010.

[PR06] C. Peikert and A. Rosen. Efficient collision-resistant hashing from worst-case assumptions on cyclic lattices. In *TCC*, pages 145–166. 2006.

[PVW08] C. Peikert, V. Vaikuntanathan, and B. Waters. A framework for efficient and composable oblivious transfer. In *CRYPTO*, pages 554–571. 2008.

[PW08] C. Peikert and B. Waters. Lossy trapdoor functions and their applications. In *STOC*, pages 187–196. 2008.

[Reg02] O. Regev. Quantum computation and lattice problems. *SIAM J. Comput.*, 33(3):738–760, 2004. Preliminary version in FOCS 2002.

[Reg03] O. Regev. New lattice-based cryptographic constructions. *J. ACM*, 51(6):899–942, 2004. Preliminary version in STOC 2003.

[Reg05] O. Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6):1–40, 2009. Preliminary version in STOC 2005.

[Reg10a] O. Regev. The learning with errors problem. In *Proc. of 25th IEEE Annual Conference on Computational Complexity (CCC)*, pages 191–204. 2010.

[Reg10b] O. Regev. On the complexity of lattice problems with polynomial approximation factors. In P. Nguyen and B. Vallée, editors, *The LLL Algorithm: Survey and Applications*. Springer-Verlag, New York, 2010.

[SJ89] A. Sinclair and M. Jerrum. Approximate counting, uniform generation and rapidly mixing Markov chains. *Inform. and Comput.*, 82(1):93–133, 1989.