New Lattice Based Cryptographic Constructions

Oded Regev *

Abstract

We introduce the use of methods from harmonic analysis as an integral part of a lattice based construction. The tools we develop provide an elegant description of certain Gaussian distributions around lattice points. Our results include two cryptographic constructions which are based on the worst-case hardness of the unique shortest vector problem. The main result is a new public key cryptosystem whose security guarantee is considerably stronger than previous results ($O(n^{1.5})$ instead of $O(n^7)$). This provides the first alternative to Ajtai and Dwork's original 1996 cryptosystem. Our second result is a collision resistant hash function which, apart from improving the security in terms of the unique shortest vector problem, is also the first example of an analysis which is not based on Ajtai's iterative step. Surprisingly, the two results are derived from the same tool which presents two indistinguishable distributions on the segment $[0, 1)$. It seems that this tool can have further applications and as an example we mention how it can be used to solve an open problem related to quantum computation.

1 Introduction

Cryptographic constructions based on lattices have attracted considerable interest in recent years. The main reason is that, unlike many other cryptographic constructions, lattice based constructions can be based on the *worst case* hardness of a problem. That is, breaking them would imply a solution to *any* instance of a certain lattice problem. In this paper we will be interested in the unique shortest vector problem (uSVP), a lattice problem which is believed to be hard: we are asked to find the shortest vector in an n-dimensional lattice with the promise that it is shorter by a factor of n^c than all other non-parallel vectors. Hence, the problem becomes harder as c decreases. The results in this field can be divided into two types. The first includes public key cryptosystems and the second includes collision resistant hash functions.

The only previously known public key cryptosystem based on a worst case lattice problem is the one due to Ajtai and Dwork [2] which appeared in 1996. They presented a public key cryptosystem based on the worst case hardness of $O(n^8)$-uSVP. Then, in [8], Goldreich, Goldwasser and Halevi showed how to eliminate decryption errors that existed in the original scheme. They also improved the security to $O(n^7)$-uSVP. Although there are other lattice based cryptosystems (see, e.g., [9, 10, 12]), none of them is based on the worst case hardness of a lattice problem. Our main result is a new public key cryptosystem whose security is based on $O(n^{1.5})$-uSVP.

In [1], Ajtai presented a one-way function based on the worst case hardness of several lattice problems. In terms of the uSVP, it was based on the hardness of $O(n^c)$-uSVP. The constant c was not explicitly specified but later it was noted to be $c = 19$ [4]. In [7], it was shown that under the same assumptions one can obtain collision resistant hash functions. These are stronger primitives than one-way functions with many uses in cryptography. Cai and Nerurkar [5] improved

*Institute for Advanced Study, Princeton, NJ. E-Mail: odedr@ias.edu. Research supported by NSF grant CCR-9987845.

the exponent to $c = 9 + \epsilon$ and later, by providing an improved analysis, Cai [4] obtained $c = 4 + \epsilon$. These papers also showed how to base the security of the hash function on other lattice problem (e.g., GapSVP and GapSIVP). Related results by Micciancio include [14] where the security is based on the covering radius problem and [13] where cyclic lattices are considered. Roughly speaking, all of these results are based on variations of a method known as Ajtai's iterative step.

Our contribution

The main contribution of this paper is the introduction of methods from harmonic analysis as an integral part of a lattice based construction. These methods were previously used indirectly through transference theorems, i.e., theorems that relate properties of a lattice and its dual (see, e.g., [4]). Our constructions are the first to make direct use of the methods from harmonic analysis.

Our main tool is a reduction from the $O(n^{1.5})$-uSVP to the problem of distinguishing between two types of distributions on the segment $[0, 1)$. We believe that this tool will find other uses in the future. A key ingredient of the main tool is the reduction from n-dimensions to one dimension. We believe that new method is another important contribution of this paper.

Using the main tool, we present three results. The main one is a new public key cryptosystem which is based on the hardness of $O(n^{1.5})$-uSVP. This is a major improvement to the 1996 cryptosystem by Ajtai and Dwork. Its description is surprising in that it essentially consists only of numbers modulo some large number N. Our second result is a collision resistant hash function whose security is based on the $O(n^{1.5})$-uSVP. In terms of the uSVP, this improves all the previous results mentioned above. More importantly though, ours is the first collision resistant hash function whose analysis is not based on Ajtai's iterative step. This might explain the improved security guarantee. The third result is related to an open question is quantum computation and will be discussed in Appendix C.

Intuitive overview

In the following we provide an informal overview of the results in this paper. Many of the details are omitted for the sake of clarity.

Main tool: Our main tool is a reduction from $O(n^{1.5})$-uSVP to the problem of distinguishing between two types of distributions on $[0, 1)$. One distribution is the uniform distribution U while the other T_h is concentrated around integer multiples of $1/h$ for some *unknown* large integer $h = 2^{O(n^2)}$ (we could easily distinguish between the two if we knew h). The sharpness of the concentration depends on the factor of the uSVP problem. For example, $O(n^{1.5})$ translates to a concentration of around $1/n$, that is, the difference between two adjacent peaks is roughly n times the width of a peak (see Figure 1). This tool is constructed in several steps. The first step involves a reduction from uSVP to a decision problem. Assume that the shortest vector is $\sum_{i=1}^{n} a_i v_i$ where $a_i \in \mathbb{Z}$ and $v_1, \ldots, v_n$ is a basis of the lattice. The decision problem asks whether $p \mid a_i$ where p is some prime number which we choose to be more than $n^{1.5}$. The reduction is a Cook reduction and the idea is to make the lattice sparser and sparser without losing the shortest vector. At the end, the lattice is so sparse that we can easily find the shortest vector. For example, when $p \mid a_i$ we can replace v_i with $p \cdot v_i$ without losing the shortest vector. The actual proof is slightly more involved as we have to handle cases where $p \nmid a_i$.

In the second step we reduce the decision problem to a problem of distinguishing between two n-dimensional distributions. Namely, one distribution is uniform and the other is a 'wavy' distribution. We begin by developing a few lemmas based on results from harmonic analysis. Essentially, these results say that if we choose a 'random' lattice point from the dual L^* of a lattice and perturb it with a Gaussian of radius 1 then the distribution obtained can be closely approximated by a function that depends only on points in L (the primal lattice) that are within distance $\sqrt{n}$ of the origin. We apply these results for two types of lattices L. The first is a lattice L where all nonzero vectors are of length more than $\sqrt{n}$. Here we get that the distribution around points of L^* is determined only by the origin of the primal lattice and is therefore very close to being uniform. The second type is a lattice with one short vector u of length (say) $1/n$ and all other non-parallel vectors of length more than $\sqrt{n}$. The distribution that we obtain here is almost uniform on $n-1$ dimensional hyperplanes orthogonal to u. In the direction of u the distribution has peaks of distance n such that the width of each peak is 1. The way we use these results is the following. Recall that we are given an $n^{1.5}$-unique lattice and we should decide whether $p \mid a_i$. We do that by first scaling the lattice so that the length of the shortest vector is $1/n$ and therefore all non-parallel vectors are of length more than $n^{1.5}/n = \sqrt{n}$. We then multiply v_i by p. If $p \mid a_i$ then the shortest vector remains and therefore if we take the distribution in the dual lattice we get a wavy distribution as described above. Otherwise, if $p \nmid a_i$, the shortest vector disappears and since $p > n^{1.5}$ the resulting lattice has no vectors shorter than $\sqrt{n}$. Therefore, the distribution obtained in the dual is very close to uniform.

The third and final step consists of 'projecting' the n-dimensional distributions described above to a one-dimensional distribution. Naïvely, one can choose a point according to the n-dimensional distribution and project it down to a line. However, this would ruin the original distribution. We would like to project down to a line but only from tiny areas around the line. This would guarantee that the original distribution is preserved. This, however, presents a new difficulty: how can one guarantee that a randomly selected point according to the distribution in $\mathbb{R}^n$ falls close to the line? We solve this by using the fact that the distribution is periodic on the lattice. Hence it is enough to consider the distribution on the fundamental parallelepiped of the lattice. Now we can draw a line that passes through the parallelepiped many times and that is therefore 'dense' in the n-dimensional space inside the parallelepiped (see Figure 2). Projecting the two n-dimensional distributions above will produce either the uniform distribution U or the wavy distribution T_h for some h. This completes the description of the main tool.

Public key cryptosystem: Let N be some large integer. The private key consists of a single integer h chosen randomly in the range (say) $[\sqrt{N}, 2\sqrt{N})$. The public key consists of $m = O(\log N)$ numbers $a_1, \ldots, a_m$ in $\{0, 1, \ldots, N-1\}$ which are 'close' to integer multiples of N/h (notice that h doesn't necessarily divide N). We also include in the public key an index $i_0 \in [m]$ such that a_{i_0} is close to an *odd* multiple of N/h. We encrypt one bit at a time. An encryption of the bit 0 is the sum of a random subset of $\{a_1, \ldots, a_m\}$ modulo N. An encryption of the bit 1 is similar but we add $\lfloor a_{i_0}/2 \rfloor$ to the result. On receiving an encrypted word w we consider its remainder on division by N/h. If it is small, we decrypt 0 and otherwise we decrypt 1. Notice that since $a_1, \ldots, a_m$ are all close to integer multiples of N/h any encryption of 0 is also close to a multiple of N/h and the decryption is correct. Similarly, since $\lfloor a_{i_0}/2 \rfloor$ is far from a multiple of N/h, encryptions of 1 are also far from multiples of N/h and the decryption is 1.

The following is a rough description of how we establish the security of the cryptosystem. Assume that there exists a distinguisher $\mathcal{A}$ that given the public key can distinguish encryptions of 0 from encryptions of 1. In other words, the difference between the acceptance probabilities p_0 on encryptions of 0 and the acceptance probability p_1 on encryptions of 1 is non-negligible. Therefore, if p_u is the acceptance probability on random words w, it must be the case that either $|p_u - p_0|$ or $|p_u - p_1|$ is non-negligible. Assume that the former case holds (the latter case is similar). Then we construct a distinguisher between the distributions U and T_h. Let R be the unknown distribution on $[0, 1)$. We choose m values from R, multiply them by N and round the result. Let $a_1, \ldots, a_m$

be the result. We then estimate $\mathcal{A}$'s acceptance probability when the public key $a_1, \ldots, a_m$ (for simplicity we ignore i_0) is fixed and the word w is chosen randomly as an encryption of 0. This is done by simply calling $\mathcal{A}$ many times, each time with a new w computed according to the encryption algorithm. We also estimate $\mathcal{A}$'s acceptance probability when w is chosen uniformly from $\{0, 1, \ldots, N-1\}$ and not according to the encryption algorithm. If there is a non-negligible difference between the two estimates, we decide that R is T_h and otherwise we say that R is U. We claim that this distinguishes between U and T_h. If R is U then $a_1, \ldots, a_m$ are uniform in $\{0, 1, \ldots, N-1\}$. Because of that one can show that the distribution of encryptions of 0 is very close to the uniform distribution and therefore $\mathcal{A}$ (as well as any other algorithm) cannot have different acceptance probabilities for the two distributions. Otherwise, R is T_h and the distribution that we obtain on $a_1, \ldots, a_m$ is the same one that is used in the public key algorithm. Therefore, according to our hypothesis, $\mathcal{A}$ should have a non-negligible difference between the two cases.

Collision resistant hash function: We choose $m = O(\log N)$ random numbers $a_1, \ldots, a_m$ from $\{0, 1, \ldots, N-1\}$ and define the function as $f(b) = \sum_{i=1}^{m} b_i a_i \bmod N$ where $b \in \{0,1\}^m$. A collision finding algorithm in this case means an algorithm $\mathcal{A}$ that given random $a_1, \ldots, a_m$ finds with non-negligible probability a nonzero vector $b \in \{-1, 0, 1\}^m$ such that $\sum b_i a_i \equiv 0 (\bmod\ N)$. Using $\mathcal{A}$ we show how to build a distinguisher between U and T_h. We can assume that we have an estimate $\tilde{h}$ on h up to some small $1/poly(m)$ error. We would like to use $\tilde{h}$ to check if the distribution is concentrated around multiples of $\frac{1}{\tilde{h}}$. But notice that the difference between $i \cdot \frac{1}{h}$ and $i \cdot \frac{1}{\tilde{h}}$ becomes huge when i is close to h (recall that h is roughly $\sqrt{N}$ which is exponential in m). The idea is to use the collision finding algorithm to create from T_h a distribution which is also concentrated around the peaks $i \cdot \frac{1}{h}$ but only for $i \leq m$.

We sample m values $x_1, \ldots, x_m$ from the unknown distribution R. Simply calling $\mathcal{A}$ with $\lfloor N \cdot x_1 \rfloor, \ldots, \lfloor N \cdot x_m \rfloor$ would yield no information as we will just have a subset S such that $\sum_{i \in S} x_i \bmod 1$ is very close to zero. So we add small perturbations $y_1, \ldots, y_m$ chosen uniformly in $[0, \frac{1}{\tilde{h}})$ to each $x_1, \ldots, x_m$ respectively. We denote the result by $z_1, \ldots, z_m$. Now we call $\mathcal{A}$ with $\lfloor N \cdot z_1 \rfloor, \ldots, \lfloor N \cdot z_m \rfloor$ and we get a subset S such that $\sum_{i \in S} z_i \bmod 1$ is very close to zero. For simplicity assume that it is exactly zero. We then check if $\sum_{i \in S} x_i \bmod 1 = -\sum_{i \in S} y_i \bmod 1$ is close to an integer multiple of $\frac{1}{\tilde{h}}$. If R is the uniform distribution on $[0, 1)$ then conditioned on any values of $z_1, \ldots, z_m$ the distribution of $x_1, \ldots, x_m$ is still uniform in $[0, \frac{1}{\tilde{h}})$ and hence $\sum_{i \in S} x_i$ is not close to an integer multiple of $\frac{1}{\tilde{h}}$. If R is T_h then conditioned on any values of $z_1, \ldots, z_m$, the y_i's are distributed around one or two peaks of T_h. Therefore, $\sum_{i \in S} y_i \bmod 1$ is close to a multiple of $\frac{1}{h}$. Moreover, since the x_i's are at most $\frac{1}{\tilde{h}}$, their sum is at most $\frac{m}{\tilde{h}}$. Since the estimate $\tilde{h}$ satisfies that for $1 \leq i \leq m$, $\frac{i}{h}$ is very close to $\frac{i}{\tilde{h}}$, the distinguisher sees that $\sum_{i \in S} x_i$ is close to a multiple of $\frac{1}{\tilde{h}}$, as required.

One last issue that we have to address is that $\mathcal{A}$ might not find collisions on inputs of the form $\lfloor N \cdot z_1 \rfloor, \ldots, \lfloor N \cdot z_m \rfloor$ when R is not the uniform distribution. That is because our assumption was that $\mathcal{A}$ finds collisions on inputs chosen uniformly. But note that if $\mathcal{A}$ does not find collisions we know that R has to be T_h and hence we can still distinguish between U and T_h.

Outline

In Section 2 we list several definitions and some properties of lattices that will be needed in this paper (for an introduction to lattices see [15]). After defining several distributions in Section 2.1 we present the two cryptographic constructions in Section 3. The main tool of the paper is developed in Section 4. The analysis of the public key cryptosystem is in Appendix A and that of the hash function is in Appendix B. In Appendix C we present a solution to an open problem related to quantum computation. Several technical claims appear in Appendix D and many proofs are in

Appendix E.

2 Preliminaries

A lattice in $\mathbb{R}^n$ is defined as the set of all integer combinations of n linearly independent vectors. This set of vectors is known as a basis of the lattice and is not unique. Given a basis $(v_1, \ldots, v_n)$ of a lattice L, the fundamental parallelepiped is defined as

$$\mathcal{P}(v_1, \ldots, v_n) = \{\sum_{i=1}^{n} x_i v_i | x_i \in [0, 1)\}.$$

When the basis is clear from the context we will use the notation $\mathcal{P}(L)$ instead of $\mathcal{P}(v_1, \ldots, v_n)$. Note that a lattice has a different fundamental parallelepiped for each possible basis. We denote by $d(L)$ the volume of the fundamental parallelepiped of L or equivalently, the determinant of the matrix B whose columns are the basis vectors of the lattice. The point $x \in \mathbb{R}^n$ reduced modulo the parallelepiped $\mathcal{P}(v_1, \ldots, v_n)$ is the unique point $y \in \mathcal{P}(v_1, \ldots, v_n)$ such that $y - x$ is an integer combination of $v_1, \ldots, v_n$ (see, e.g., [12]). The dual of a lattice L in $\mathbb{R}^n$, denoted L^*, is the set of all vectors $y \in \mathbb{R}^n$ such that $\langle x, y\rangle \in \mathbb{Z}$ for all vectors $x \in L$. Similarly, given a basis $(v_1, \ldots, v_n)$ of lattice, we define the dual basis as the set of vectors $(v_1^*, \ldots, v_n^*)$ such that $\langle v_i, v_j^*\rangle = \delta_{ij}$ for all $i, j \in [n]$. Note that if $B = (v_1, \ldots, v_n)$ is the $n \times n$ matrix whose columns are the basis vectors then $(B^T)^{-1}$ contains the dual basis as its columns. From this it follows that $d(L^*) = 1/d(L)$.

We say that a lattice is *unique* if its shortest vector is strictly shorter than all other non-parallel vectors. Moreover, a lattice is $f(n)$*-unique* if the shortest vector is shorter by a factor of at least $f(n)$ from all non-parallel vectors. In the shortest vector problem we are interested in finding the shortest vector in a lattice. In this paper we will be interested in the $f(n)$-unique shortest vector problem ($f(n)$-uSVP) where in addition, we are promised that the lattice is $f(n)$-unique. Let $\lambda(L)$ denote the length of the shortest nonzero vector in the lattice L. We also denote the shortest vector (or one of the shortest vectors) by $\tau(L)$. Most of the lattices that appear is this paper are unique lattices and in these cases $\tau(L)$ is unique up to sign.

One particularly useful type of basis is an LLL reduced basis. Such a basis can be found in polynomial time [11]. Hence, we will often assume without loss of generality that lattices are given by an LLL reduced basis. The properties of LLL reduced bases that we use are summarized in Claim D.6.

We define a negligible amount as an amount which is asymptotically smaller than n^{-c} for any constant $c > 0$. The parameter n will indicate the input size. Similarly, a non-negligible amount is one which is at least n^{-c} for some $c > 0$. Finally, exponentially small means an expression that is at most $2^{-\Omega(n)}$. We say that an algorithm $\mathcal{A}$ with oracle access is a distinguisher between two distributions if its acceptance probability when the oracle outputs samples of the first distribution and its acceptance probability when the oracle outputs samples of the second distribution differ by a non-negligible amount. Note that the notion of acceptance is used for convenience. In addition, an algorithm $\mathcal{A}$ is said to distinguish between the distribution T and the set of distribution $\mathcal{T}$ if for any distribution $T' \in \mathcal{T}$, $\mathcal{A}$ distinguishes between T and T'.

For two continuous random variables X and Y having values in $[0, 1)$ with density functions T_1 and T_2 respectively we define their statistical difference as

$$\Delta(X, Y) = \frac{1}{2}\int_0^1 |T_1(r) - T_2(r)| dr.$$

A similar definition holds for discrete random variables. One important fact that we use is that the statistical distance cannot increase by applying a (possibly randomized) function f, i.e., $\Delta(f(X), f(Y)) \leq \Delta(X, Y)$. In particular, this implies that the acceptance probability of any algorithm on inputs from X differs from its acceptance probability on inputs from Y by at most $\Delta(X, Y)$.

The set $\{1, 2, \ldots, n\}$ is denoted by $[n]$. All logarithms are of base 2 unless otherwise specified. We use δ_{ij} to denote the Kronecker delta, i.e., 1 if $i = j$ and 0 otherwise. We use $\tilde{c}$ to denote an unspecified constant. That is, whenever $\tilde{c}$ appears we can replace it with some universal constant. For example, the expression $\tilde{c} + 7 = \tilde{c}$ is true because we can substitute 1 and 8 for the constants. Other constants will be denoted by c with a letter as the subscript, e.g., c_{h}.

For two real numbers $x, y > 0$ we define $x \bmod y$ as $x - \lfloor x/y \rfloor y$. For $x \in \mathbb{R}$ we define $\lfloor x \rceil$ as the integer nearest to x or, in case two such integers exist, the smaller of the two. We also use the notation $\operatorname{frc}(x) := |x - \lfloor x \rceil|$, i.e., the distance of a real x to the nearest integer. Notice that for all $x, y \in \mathbb{R}$, $0 \leq \operatorname{frc}(x) \leq \frac{1}{2}$, $\operatorname{frc}(x) \leq |x|$ and $\operatorname{frc}(x + y) \leq \operatorname{frc}(x) + \operatorname{frc}(y)$.

Recall that the *normal distribution* with mean 0 and variance σ^2 is the distribution on $\mathbb{R}$ given by the density function $\frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{1}{2}(\frac{x}{\sigma})^2}$. Also recall that the sum of two normal variables with mean 0 and variances σ_1^2 and σ_2^2 is a normal variable with mean 0 and variance $\sigma_1^2 + \sigma_2^2$. The *Gaussian* distribution is a distribution on $\mathbb{R}^n$ obtained by taking n independent identically distributed normal random variables as the coordinates. A simple tail bound on the normal distribution appears in Claim D.1.

For clarity, we present some of our reductions in a model which allows operations on real numbers. It is possible to modify them in a straight forward way so that they operate in a model that approximates real numbers up to an error of 2^{-n^c} for arbitrary large constant c in time polynomial in n. Therefore, if we say that two continuous distributions on $[0, 1)$ are indistinguishable (in the real model) then for any $c > 0$ discretizing the distributions up to error 2^{-n^c} for any c yields two indistinguishable distributions.

2.1 Several Distributions

We define several useful distributions on the segment $[0, 1)$. The distribution U is simply the uniform distribution with the density function $U(r) = 1$. For $\beta \in \mathbb{R}^+$ the distribution Q_β is a normal distribution with mean 0 and variance $\frac{\beta}{2\pi}$ reduced modulo 1 (i.e., a periodization of the normal distribution):

$$Q_\beta(r) = \sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{\beta}} e^{-\frac{\pi}{\beta}(r-k)^2}.$$

Clearly, one can efficiently sample from Q_β by sampling a normal variable and reducing the result modulo 1. Another distribution is $T_{h,\beta}$ where $h \in \mathbb{N}$ and $\beta \in \mathbb{R}^+$ (see Figure 1). Its density function is defined as

$$T_{h,\beta}(r) = Q_\beta(rh \bmod 1) = \sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{\beta}} e^{-\frac{\pi}{\beta}(rh-k)^2}.$$

By adding a normalization factor we can extend the definition of $T_{h,\beta}$ to non-integer h. So in general,

$$T_{h,\beta}(r) = \frac{1}{\int_0^1 Q_\beta(xh \bmod 1) dx} Q_\beta(rh \bmod 1).$$

For a real $h > 0$, choosing a value $z \in [0, 1)$ according to $T_{h,\beta}$ can be done as follows. First choose a value $x \in \{0, 1, \ldots, \lceil h \rceil - 1\}$ and then choose a value y according to Q_β. If $\frac{x+y}{h} < 1$ then return it as the result. Otherwise, repeat the process again. It is easy to see that the distribution obtained is indeed $T_{h,\beta}$ and that the process is efficient for (say) $h \geq 1$.

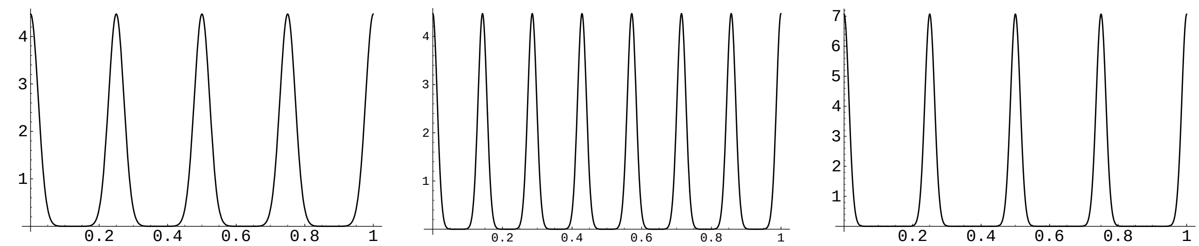

Figure 1: $T_{4,0.05}$, $T_{7,0.05}$ and $T_{4,0.02}$

We also define the following set of distributions:

$$\mathcal{T}_{g(n)} := \{T_{h,\beta} \mid h \in \mathbb{N},\ h \leq 2^{c_{\mathsf{h}} n^2},\ \beta \in [\frac{n}{(g(n))^2}, 4\frac{n}{(g(n))^2})\}$$

where c_{h} is a constant specified is Lemma 4.9.

3 Cryptographic Constructions

For a security parameter n, let N be $2^{c_{\mathsf{N}} n^2}$ and let m be $c_{\mathsf{m}} n^2$ where c_{N} and c_{m} are two constants which will be specified later. Let $\gamma(n) = \omega(n\sqrt{\log n})$, i.e., any function that satisfies $\frac{\gamma(n)}{n\sqrt{\log n}} \to \infty$ as n goes to infinity. The smaller the function, the better the security guarantee becomes. For concreteness, one can choose $\gamma(n) = n \log n$. We also assume that $\gamma(n) \leq n^{c_\gamma}$ for some constant $c_\gamma > 0$.

Public Key Encryption

- **Private key:** Let $H = \{h \in [\sqrt{N}, 2\sqrt{N}) \mid \mathrm{frc}(h) < \frac{1}{16m}\}$. Choose $h \in H$ uniformly at random. Let d denote $\frac{N}{h}$. The private key is the number h.
- **Public Key:** Choose $\beta \in [4\frac{1}{(\gamma(n))^2}, 8\frac{1}{(\gamma(n))^2})$ uniformly at random. We choose m values $z_1, \ldots, z_m$ from $T_{h,\beta}$ by choosing $x_1, \ldots, x_m$ and $y_1, \ldots, y_m$ as described in Section 2.1. Let i_0 be an index such that x_{i_0} is odd (such an i_0 exists with probability exponentially close to 1). For $i \in [m]$, let a_i denote $\lfloor N \cdot z_i \rfloor$. The public key is $(a_1, \ldots, a_m, i_0)$.
- **Encryption:** In order to encrypt a bit we choose a random subset S of $[m]$. The encryption is $\sum_{i \in S} a_i \bmod N$ if the bit is 0 and $\sum_{i \in S} a_i + \lfloor \frac{a_{i_0}}{2} \rfloor \bmod N$ if the bit is 1.
- **Decryption:** On receiving $w \in \{0, \ldots, N-1\}$ we decrypt 0 if $\mathrm{frc}(\frac{w}{d}) < \frac{1}{4}$ and 1 otherwise.

Collision Resistant Hash Function

- Choose m numbers $a_1, \ldots, a_m$ uniformly in $\{0, 1, \ldots, N-1\}$. The function $f : \{0,1\}^m \to \{0, 1, \ldots, N-1\}$ is defined as:

$$f(b) = \sum_{i=1}^{m} b_i a_i \bmod N.$$

Notice that if $c_m > c_N$ then f indeed compresses the size of the input and collisions are guaranteed to exist.

4 Main Tool

In this section we present a reduction from $g(n)$-uSVP to the problem of distinguishing between two types of distributions on $[0, 1)$.

Theorem 4.1 *Let $g(n)$ be any function such that $4\sqrt{n} \leq g(n) \leq poly(n)$ and let c_{h} the constant specified in Lemma 4.9. If there exists a distinguisher between U and $\mathcal{T}_{g(n)}$ then there exists a solution to $g(n)$-uSVP.*

Proof: Let $p(n)$ be a prime larger than $g(n)$ and at most (say) $2g(n)$. We can now apply Lemmas 4.2, 4.8 and 4.9 in order to obtain the theorem. ∎

4.1 Reduction to a Decision Problem

We reduce the SVP to the following decision problem:

Decision SVP with parameter p ($dSVP_p$)

- **Input:** An arbitrary basis $(v_1, \ldots, v_n)$ of a unique lattice L and a number α such that $\lambda(L) < \alpha \leq 2\lambda(L)$ and let $\tau(L) = \sum_{i=1}^{n} a_i v_i$ be the coefficients of the shortest vector.
- **Output:** YES if p divides a_1, NO otherwise.

Lemma 4.2 *Let $p = p(n) > 2$ be a prime number which is at most polynomial in n.*[1] *There exists a reduction from finding the shortest vector in a unique lattice L to $dSVP_p$.*[2] *Moreover, if L is an $f(n)$-unique lattice then all the calls to $dSVP$ are also with an $f(n)$-unique lattice.*

Although not used in this paper, the following is an immediate corollary of the above lemma and might be of independent interest. Basically, it is a reduction from the search SVP to the decision SVP for unique lattices. It is still an open question whether a similar result holds for SVP on general lattices.

Corollary 4.3 *For any prime $p = p(n) < poly(n)$ larger than 2 and any $f(n) \geq 1$, finding the shortest vector in an $p(n)f(n)$-unique lattice can be reduced to the following gap problem: given d and an $f(n)$-unique lattice, decide whether the length of the shortest vector is at most d or more than $\sqrt{p(n)} \cdot d$.*

4.2 Gaussian Distributions on Lattices

Let B_n denote the Euclidean unit ball and define $\rho(A)$ as $\sum_{x \in A} e^{-\pi\|x\|^2}$. The following lemma says that in any lattice L the contribution to $\rho(L)$ from points of distance more than $\sqrt{n}$ is negligible.

Lemma 4.4 ([3], Lemma 1.5(i) with $c = 1$) *For any lattice L, $\rho(L - \sqrt{n}B_n) < 2^{-\Omega(n)}\rho(L)$.*

A simple corollary of this is that $\rho(L) < \rho(L \cap \sqrt{n}B_n)/(1 - 2^{-\Omega(n)})$. We will also use the following formulation of the Poisson summation formula:

Lemma 4.5 ([3], Lemma 1.1(i) with $a = \pi$, $b = 1$, $y = 0$) *For any lattice L and any vector $y \in \mathbb{R}^n$, $\rho(L^* + y) = d(L) \cdot \sum_{x \in L} e^{2\pi i \langle x, y \rangle} \rho(\{x\})$.*

[1] The result holds for the case $p = 2$ as well with some technical differences.

[2] One can guarantee the uniqueness of the shortest vector in any lattice by adding tiny perturbations to the basis vectors. Therefore, the assumption that L is unique can be avoided.

For a given lattice L, we consider the distribution obtained by sampling a Gaussian with standard deviation $1/\sqrt{2\pi}$ centered around the origin and reducing it modulo the fundamental parallelepiped $\mathcal{P}(L^*)$. Equivalently, we consider the following density function defined on $\mathcal{P}(L^*)$:

$$D_{L^*}(x) = \rho(L^* + x).$$

Intuitively, we can think of D_{L^*} as taking Gaussian distributions around 'all' points of L^*. Since this distribution is periodic in $\mathbb{R}^n$ with period $\mathcal{P}(L^*)$, we simplify the analysis by choosing D_{L^*} to be a restriction of the distribution to $\mathcal{P}(L^*)$. In this section we present good approximations to D_{L^*} for two types of lattices L.

Lemma 4.6 *Let L be a lattice in which all non-zero vectors are of length more than $\sqrt{n}$ and let $U_{L^*}(x) = \frac{1}{d(L^*)} = d(L)$ be the uniform density function on $\mathcal{P}(L^*)$. Then, $\Delta(D_{L^*}, U_{L^*}) < 2^{-\Omega(n)}$.*

For any vector $v \in L$ define the density function $T_{L^*,v}$ on $\mathcal{P}(L^*)$ as

$$T_{L^*,v}(x) = \frac{d(L)}{\|v\|} \sum_{k\in\mathbb{Z}} e^{-\pi(\frac{k+\langle v,x\rangle}{\|v\|})^2}.$$

According to the Claim D.4 it is indeed a density function.

Lemma 4.7 *Let L be a lattice with a shortest vector u in which all vectors not parallel to u are of length more than $\sqrt{n}$. Then, $\Delta(D_{L^*}, T_{L^*,u}) < 2^{-\Omega(n)}(1 + \frac{1}{\|u\|})$. In particular, if $\|u\| \geq \frac{1}{n^c}$ for some $c > 0$ then $\Delta(D_{L^*}, T_{L^*,u}) < 2^{-\Omega(n)}$.*

4.3 Two Indistinguishable Distributions

Lemma 4.8 *Let $g(n) < p(n)$ be such that $p(n)$ is a prime and both are at most polynomial in n. Solving $dSVP_{p(n)}$ on $g(n)$-unique lattices can be reduced to the problem of distinguishing between U_{L^*} and $T_{L^*,\tau(L)}$ where L is given as an LLL reduced basis and $\lambda(L) \in [\frac{\sqrt{n}}{g(n)}, \frac{2\sqrt{n}}{g(n)})$.*

4.4 One Dimensional Distributions

Lemma 4.9 *There exists a constant c_{h} such that for any $g(n) \geq 4\sqrt{n}$, $g(n) \leq poly(n)$, the problem of distinguishing between U_{L^*} and $T_{L^*,\tau(L)}$ for a lattice L given as an LLL reduced basis for which $\lambda(L) \in [\frac{\sqrt{n}}{g(n)}, \frac{2\sqrt{n}}{g(n)})$ can be reduced to the problem of distinguishing between U and $\mathcal{T}_{g(n)}$.*

Proof: Let $v_1, \ldots, v_n$ denote the LLL reduced basis of L and let $v_1^*, \ldots, v_n^*$ be the dual basis of L^*, i.e., a basis of L^* such that $\langle v_i, v_j^* \rangle = \delta_{ij}$. For some large integer K to be chosen later, consider a function f which maps a vector $v = \sum_{i=1}^n a_i v_i^*$ in $\mathcal{P}(L^*)$ to $\frac{\lfloor Ka_1 \rfloor}{K} + \frac{\lfloor Ka_2 \rfloor}{K^2} + \ldots + \frac{\lfloor Ka_{n-1} \rfloor}{K^{n-1}} + \frac{a_n}{K^n} \in [0,1)$. An equivalent way to describe f is the following. For a real $r \in [0,1)$ let $r_1, \ldots, r_{n-1} \in \{0, \frac{1}{K}, \ldots, \frac{K-1}{K}\}$ and $r_n \in [0,1)$ be the unique numbers such that $r = r_1 + \frac{1}{K} r_2 + \ldots + \frac{1}{K^{n-2}} r_{n-1} + \frac{1}{K^{n-1}} r_n$. The set of points that are mapped to r is given by

$$S(r) := \{\sum_{i=1}^n a_i v_i^* \mid \forall i \in [n-1] \; a_i \in [r_i, r_i + \frac{1}{K}] \text{ and } a_n = r_n\}.$$

Hence, $S(r)$ is an $n-1$ dimensional parallelepiped whose diameter is at most $\frac{1}{K}\sum_{i=1}^{n-1} \|v_i^*\|$. Let w denote the vector $v_1^* + K v_2^* + \ldots + K^{n-1} v_n^*$. Then, it is easy to see that for any $r \in [0,1)$ the point

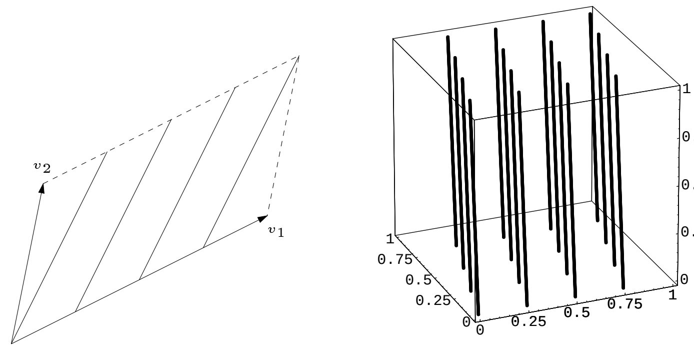

Figure 2: The line connecting the origin to w with $K = 4$ in two dimensions with $\mathcal{P}(v_1, v_2)$ and in three dimensions with the unit cube.

rw reduced modulo $\mathcal{P}(L^*)$ is contained in $S(r)$. The line connecting the origin with w reduced modulo $\mathcal{P}(L^*)$ goes through the parallelepiped K^{n-1} times. The mapping f essentially takes each point in $\mathcal{P}(L^*)$ to a nearby point on the line (see Figure 2).

The reduction works by sampling a point from the given distribution on $\mathcal{P}(L^*)$ and applying f, thereby obtaining a distribution on $[0, 1)$. Notice that f can be computed efficiently. Clearly, by starting from a uniform distribution on $\mathcal{P}(L^*)$ we obtain the uniform distribution on $[0, 1)$. Hence, it is enough to consider $T_{L^*,\tau(L)}$. The distribution that we get on $[0, 1)$ is given by:

$$T_1(r) := \frac{d(L^*)}{\text{vol}(S(r))} \int_{S(r)} T_{L^*,\tau(L)}(x)dx$$

which is $d(L^*)$ times the average of $T_{L^*,\tau(L)}$ over $S(r)$. We claim that by choosing K to be large enough this average is very close to its value in $rw \in S(r)$. More formally, we claim that $T_1(r)$ is exponentially close to

$$T_{|\langle\tau(L),w\rangle|,\lambda(L)^2}(r) = \frac{1}{\lambda(L)} \sum_{k\in\mathbb{Z}} e^{-\pi(\frac{k+r\langle\tau(L),w\rangle}{\lambda(L)})^2} = d(L^*)T_{L^*,\tau(L)}(rw)$$

where in the first equality we used the fact that $\langle\tau(L), w\rangle$ is integer and that the function does not change if we change the sign of $\langle\tau(L), w\rangle$.

By using the mean value theorem we get that for any $r \in [0, 1)$ the difference between the maximum and the minimum values of $T_{L^*,\tau(L)}$ over $S(r)$ is at most:

$$\text{diam}(S(x)) \cdot \max_x \frac{d}{dx}\left(\frac{d(L)}{\lambda(L)} \sum_{k\in\mathbb{Z}} e^{-\pi(\frac{k+\lambda(L)x}{\lambda(L)})^2}\right) \leq \tilde{c} \cdot \frac{1}{K} \sum_{i=1}^{n-1} \|v_i^*\| \cdot \frac{d(L)}{\lambda(L)}$$

where the inequality is due to Claim D.3 and the assumption that $\lambda(L) \leq \frac{2\sqrt{n}}{g(n)} \leq \frac{1}{2}$. Hence, using Claim D.6,

$$\begin{aligned} \forall r \in [0,1) \quad \left|T_1(r) - T_{|\langle\tau(L),w\rangle|,\lambda(L)^2}(r)\right| &\leq d(L^*) \cdot \tilde{c} \cdot \frac{1}{K} \sum_{i=1}^{n-1} \|v_i^*\| \cdot \frac{d(L)}{\lambda(L)} = \tilde{c} \cdot \frac{1}{K} \sum_{i=1}^{n-1} \|v_i^*\| \cdot \frac{1}{\lambda(L)} \\ &\leq \tilde{c} \cdot \frac{1}{K} \cdot n \cdot \frac{\sqrt{n}}{\lambda(L)} \cdot 2^{2n} \cdot \frac{1}{\lambda(L)} \leq \tilde{c} \cdot \frac{1}{K} \cdot 2^{2n} \cdot poly(n) \end{aligned}$$

and by choosing $K = 2^{3n}$ we get that the statistical distance between T_1 and $T_{|\langle \tau(L),w\rangle|,\lambda(L)^2}$ is exponentially small.

Recall that $w = \sum_{i=1}^{n} K^{i-1} v_i^*$ and $\tau(L) = \sum_{i=1}^{n} a_i v_i$ where all $|a_i| \leq 2^{2n}$. Since $\langle v_i, v_j^* \rangle = \delta_{ij}$, the inner product $\langle \tau(L), w\rangle$ is integer and its absolute value is at most $n \cdot 2^{2n} \cdot K^n \leq 2^{c_{\mathsf{h}} n^2}$ for a large enough c_{h}, as required. ■

5 Acknowledgements

I thank Irit Dinur for suggesting that I look at cryptographic constructions and Daniele Micciancio for many helpful comments on an earlier draft of this paper.

References

[1] M. Ajtai. Generating hard instances of lattice problems. In *Proc. 28th ACM Symp. on Theory of Computing*, pages 99–108, 1996.

[2] M. Ajtai and C. Dwork. A public-key cryptosystem with worst-case/average-case equivalence. In *Proc. 29th ACM Symp. on Theory of Computing*, pages 284–293, 1997.

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

[4] J.-Y. Cai. Applications of a new transference theorem to Ajtai's connection factor. In *Fourteenth Annual IEEE Conference on Computational Complexity (Atlanta, GA, 1999)*, pages 205–214. IEEE Computer Soc., Los Alamitos, CA, 1999.

[5] J.-Y. Cai and A. P. Nerurkar. An improved worst-case to average-case connection for lattice problems (extended abstract). In *38th Annual Symposium on Foundations of Computer Science*, pages 468–477, 1997.

[6] M. Ettinger and P. Høyer. On quantum algorithms for noncommutative hidden subgroups. *Advances in Applied Mathematics*, 25(3):239–251, 2000.

[7] O. Goldreich, S. Goldwasser, and S. Halevi. Collision-free hashing from lattice problems. In *ECCCTR: Electronic Colloquium on Computational Complexity, technical reports*, 1996.

[8] O. Goldreich, S. Goldwasser, and S. Halevi. Eliminating decryption errors in the Ajtai-Dwork cryptosystem. *Lecture Notes in Computer Science*, 1294:105, 1997.

[9] O. Goldreich, S. Goldwasser, and S. Halevi. Public-key cryptosystems from lattice reduction problems. In *Advances in cryptology (CRYPTO) '97 (Santa Barbara, CA, 1997)*, volume 1294 of *Lecture Notes in Comput. Sci.*, pages 112–131. Springer, 1997.

[10] J. Hoffstein, J. Pipher, and J. H. Silverman. NTRU: a ring-based public key cryptosystem. In *Algorithmic number theory (Portland, OR, 1998)*, volume 1423 of *Lecture Notes in Comput. Sci.*, pages 267–288. Springer, 1998.

[11] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261(4):515–534, 1982.

[12] D. Micciancio. Improving lattice based cryptosystems using the hermite normal form. In *Cryptography and Lattices Conference (CaLC)*, volume 2146 of *Lecture Notes in Computer Science*, pages 126–145, Providence, Rhode Island, March 2001. Springer-Verlag.

[13] D. Micciancio. Generalized compact knapsaks, cyclic lattices, and efficient one-way functions from worst-case complexity assumptions. In *Proceedings of the 43rd Annual Symposium on Foundations of Computer Science (FOCS) 2002*, Vancouver, Canada, November 2002. To appear.

[14] D. Micciancio. Improved cryptographic hash functions with worst-case/average-case connection. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC) 2002*, pages 609–618, Montreal, Canada, 2002.

[15] D. Micciancio and S. Goldwasser. *Complexity of Lattice Problems: a cryptographic perspective*, volume 671 of *The Kluwer International Series in Engineering and Computer Science*. Kluwer Academic Publishers, Boston, Massachusetts, March 2002.

[16] O. Regev. Quantum computation and lattice problems. In *Proceedings of the 43rd Annual Symposium on Foundations of Computer Science (FOCS) 2002*, Vancouver, Canada, November 2002. To appear.

A Analysis of the Public Key Cryptosystem

Lemma A.1 (Correctness) *The probability of a decryption error is at most* $2^{-\Omega(\frac{(\gamma(n))^2}{m})}$ *plus some exponentially small terms.*

Note that the above probability is negligible since $\gamma(n) = \omega(n\sqrt{\log n})$.

Proof: First consider an encryption of the bit 0. Probabilities are taken over the choices of the private and public keys and the randomization in the encryption process. Let S denote the subset of indices which are included in the sum and let $w := \sum_{i \in S} a_i \bmod N$. Since $\sum_{i \in S} a_i \leq m \cdot N$,

$$\left| w - (\sum_{i \in S} a_i \bmod d\lfloor h \rceil) \right| \leq m \cdot |N - d\lfloor h \rceil| = m \cdot d \cdot \mathrm{frc}(h) < \frac{1}{16} d$$

and by the triangle inequality,

$$\mathrm{frc}(\frac{w}{d}) < \frac{1}{16} + \mathrm{frc}(\frac{\sum_{i \in S} a_i \bmod d\lfloor h \rceil}{d}) = \frac{1}{16} + \mathrm{frc}(\frac{\sum_{i \in S} a_i}{d}) < \frac{1}{16} + \frac{m}{d} + \mathrm{frc}(\frac{N}{d} \sum_{i \in S} z_i)$$

where the last inequality uses $|N \cdot z_i - a_i| < 1$. Notice that $\mathrm{frc}(\frac{N}{d} \sum_{i \in S} z_i) = \mathrm{frc}(\sum_{i \in S}(x_i + y_i)) = \mathrm{frc}(\sum_{i \in S} y_i)$. Hence,

$$\mathrm{frc}(\frac{w}{d}) < \frac{1}{16} + \frac{m}{d} + \mathrm{frc}(\sum_{i \in S} y_i) < \frac{1}{8} + \mathrm{frc}(\sum_{i \in S} y_i)$$

where we used the fact that d is much larger than m. With probability exponentially close to 1, all x_i's are strictly less than $\lceil h \rceil - 1$. Conditioned on that, the distribution of y_i is Q_β and the distribution of $\sum_{i \in S} y_i \bmod 1$ is $Q_{|S|\beta}$ where $|S|\beta \leq m \cdot \beta = O(\frac{m}{(\gamma(n))^2})$. Therefore, according to Claim D.1, the probability of $\mathrm{frc}(\sum_{i \in S} y_i) > \frac{1}{16}$ is at most $2^{-\Omega(\frac{\gamma(n)}{m})}$ and hence

$$\mathrm{frc}(\frac{w}{d}) < \frac{1}{8} + \frac{1}{16} \tag{1}$$

which is less than $\frac{1}{4}$, as required.

The proof for the case of an encryption of 1 is similar. By using the fact that x_{i_0} is odd and that with probability exponentially close to 1, $\mathrm{frc}(y_{i_0}) < \frac{1}{16}$ we get $\mathrm{frc}(\frac{\lfloor a_{i_0}/2 \rfloor}{d}) > \frac{1}{2} - \frac{1}{32} - \frac{1}{d}$. This, combined with (1) gives

$$\mathrm{frc}(\frac{w}{d}) > \mathrm{frc}(\frac{\lfloor a_{i_0}/2 \rfloor}{d}) - \frac{1}{8} - \frac{1}{16} > \frac{1}{4}$$

and the proof is completed. ■

Before establishing the security of the construction, let us prove a few simple claims.

Claim A.2 *For any $h \in \mathbb{N}, \beta \in \mathbb{R}$, let X, Y be two independent random variables; X is distributed uniformly over $\{0, \frac{1}{h}, \ldots, \frac{h-1}{h}\}$ and Y is normal with mean 0 and variance $\frac{\beta}{2\pi h^2}$. Then $T_{h,\beta}$ is equivalent to the distribution of the sum of X and Y reduced modulo 1.*

Proof:

$$\begin{aligned} T_{h,\beta}(r) &= Q_\beta(hr \bmod 1) = \sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{\beta}} e^{-\frac{\pi}{\beta}(hr-k)^2} = \\ &\sum_{l=0}^{h-1} \sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{\beta}} e^{-\frac{\pi}{\beta}(hr-hk-l)^2} = \sum_{l=0}^{h-1} \frac{1}{h} \sum_{k=-\infty}^{\infty} \frac{h}{\sqrt{\beta}} e^{-\frac{\pi h^2}{\beta}(r-k-\frac{l}{h})^2} \end{aligned}$$

■

Claim A.3 *For $h \in \mathbb{N}$, $T_{h,\beta} + Q_\delta$ mod $1 = T_{h,\beta+\delta h^2}$.*

Proof: According to Claim A.2, $T_{h,\beta}$ can be viewed as the sum of two random variables X and Y reduced modulo 1. Therefore, $T_{h,\beta} + Q_\delta$ mod $1 = X + Y + Q_\delta$ mod 1. But since both Y and Q_δ are normal, their sum modulo 1 is exactly $Q_{\frac{\beta}{h^2}+\delta}$ and we conclude the proof by using Claim A.2 again. ■

Definition A.4 *Given a density function X on $[0, 1)$ we define its compression by a factor $\delta \geq 1$ as the distribution on $[0, 1)$ given by*

$$\frac{1}{\int_0^1 X(\delta x \text{ mod } 1)dx} X(\delta r \text{ mod } 1).$$

We denote the result by $C_\delta(X)$.

Using the above definition, $T_{h,\beta}$ is a compression of Q_β by a factor of h. Notice that if we can sample efficiently from X then we can also sample efficiently from its compression. This is done in a way similar to that used to sample from $T_{h,\beta}$.

Claim A.5 *For any $h \in \mathbb{N}$ and $\delta \geq 1$, the compression of $T_{h,\beta}$ by a factor δ is $T_{\delta h,\beta}$.*

Proof: The proof follows directly from the definition of $T_{h,\beta}$. ■

Claim A.6 *For large enough c, when choosing $c \cdot l$ numbers $a_1, \ldots, a_{c \cdot l}$ uniformly from 0 to $2^l - 1$ the probability that the statistical distance between the uniform distribution on $\{0, \ldots, 2^l - 1\}$ and the distribution given by sums modulo 2^l of random subsets of $\{a_1, \ldots, a_{c \cdot l}\}$ is more than 2^{-l} is at most 2^{-l}.*

Proof: Let $X_{t,b}$ for $t \in \{0,\dots,2^l-1\}$, $b \in \{0,1\}^{c\cdot l} \backslash 0^{c\cdot l}$ denote the event that $\sum_{i=1}^{c\cdot l} b_i a_i \equiv t \pmod{2^l}$ where the probability is taken over the choice of $\{a_1,\dots,a_{c\cdot l}\}$. Then, $E[X_{t,b}] = 2^{-l}$ and $V[X_{t,b}] < 2^{-l}$. Hence, $E[Y_t] = \frac{2^{c\cdot l}-1}{2^l} = 2^{(c-1)\cdot l} - 2^{-l}$ where Y_t denotes $\sum_{b\in\{0,1\}^{c\cdot l}\backslash 0^{c\cdot l}} X_{t,b}$. Moreover, for $b \neq b'$, the events $X_{t,b}$ and $X_{t,b'}$ are pairwise disjoint. Therefore, $V[Y_t] < \frac{2^{c\cdot l}-1}{2^l} < 2^{(c-1)\cdot l}$. Using the Chebyshev inequality,

$$Pr\left(\left|Y_t - (2^{(c-1)\cdot l} - 2^{-l})\right| \geq 2^{(\frac{c-1}{2}+1)\cdot l}\right) \leq 2^{-2l}$$

and hence,

$$Pr\left(\left|Y_t - 2^{(c-1)\cdot l}\right| \geq 2^{(\frac{c-1}{2}+1)\cdot l} + 2^{-l}\right) \leq 2^{-2l}.$$

Using the union bound,

$$Pr\left(\exists t,\ \left|Y_t - 2^{(c-1)\cdot l}\right| \geq 2^{(\frac{c-1}{2}+1)\cdot l} + 2^{-l}\right) \leq 2^{-l}.$$

Therefore, with probability at least $1-2^{-l}$ on the choice of $\{a_1,\dots,a_{c\cdot l}\}$, the number of subsets (including the empty subset) mapped to each number t is at most $2^{(\frac{c-1}{2}+1)\cdot l} + 2^{-l} + 1 \leq 2^{(\frac{c-1}{2}+2)\cdot l}$ away from $2^{(c-1)\cdot l}$. This translates to a statistical distance of at most

$$2^{(\frac{c-1}{2}+2)\cdot l} \cdot 2^{-(c-1)\cdot l} < 2^{-l}$$

for large enough c. ■

Lemma A.7 (Security) *For $c_{\mathsf{N}} \geq 2c_{\mathsf{h}}$ and large enough c_{m}, if there exists a polynomial time algorithm $\mathcal{A}$ that distinguishes between encryptions of* 0 *and* 1 *then there exists an algorithm $\mathcal{B}$ that distinguishes between the distributions U and $\mathcal{T}_{\sqrt{n}\gamma(n)}$.*

Proof: Let p_0 be the acceptance probability of $\mathcal{A}$ on input $((a_1,\dots,a_m,i_0),w)$ where w is an encryption of 0 with the public key $(a_1,\dots,a_m,i_0)$ and the probability is taken over the choice of private and public keys and the encryption algorithm. We define p_1 similarly for encryptions of 1 and let p_u be the acceptance probability of $\mathcal{A}$ on inputs $((a_1,\dots,a_m,i_0),w)$ where $a_1,\dots,a_m,i_0$ are again chosen according to the private and public keys distribution but w is chosen uniformly from $\{0,\dots,N-1\}$. We would like to construct an $\mathcal{A}'$ that distinguishes between the case where w is an encryption of 0 and the case where w is random. According to our hypothesis, $|p_0 - p_1| \geq \frac{1}{n^c}$ for some $c > 0$. Therefore, either $|p_0 - p_u| \geq \frac{1}{2n^c}$ or $|p_1 - p_u| \geq \frac{1}{2n^c}$. In the former case $\mathcal{A}$ is itself the required distinguisher. In the latter case $\mathcal{A}$ distinguishes between the case where w is an encryption of 1 and the case where w is random. We construct $\mathcal{A}'$ as follows. On input $((a_1,\dots,a_n,i_0),w)$, $\mathcal{A}'$ calls $\mathcal{A}$ with $((a_1,\dots,a_n,i_0), w + \lfloor\frac{a_{i_0}}{2}\rfloor \bmod N)$. Notice that this maps the distribution on encryptions of 0 to the distribution on encryptions of 1 and the uniform distribution to itself. Therefore, $\mathcal{A}'$ is the required distinguisher.

Let $p_0(a_1,\dots,a_m,i_0)$ be the probability that $\mathcal{A}'$ accepts on inputs $((a_1,\dots,a_m,i_0),w)$ where the probability is taken only over the choice of w as an encryption of 0 with the fixed public key $(a_1,\dots,a_m,i_0)$. Similarly, define $p_u(a_1,\dots,a_m,i_0)$ to be the acceptance probability of $\mathcal{A}'$ where w is now chosen uniformly at random from $\{0,\dots,N-1\}$. Let Y be the set of all tuples $(a_1,\dots,a_m,i_0)$ for which $|p_0(a_1,\dots,a_m,i_0) - p_u(a_1,\dots,a_m,i_0)| \geq \frac{1}{4n^c}$. By an averaging argument we get that with probability at least $\frac{1}{4n^c}$ on the choice of $(a_1,\dots,a_m,i_0)$, $(a_1,\dots,a_m,i_0) \in Y$ for otherwise $\mathcal{A}'$ would have a gap of less than $\frac{1}{2n^c}$.

In the following we describe the distinguisher $\mathcal{B}$. We are given a distribution R which is either U or some $T_{h,\beta} \in \mathcal{T}_{\sqrt{n}\gamma(n)}$ with an integer $h \leq 2^{c_{\mathsf{h}} n^2} \leq \sqrt{N}$ and a real $\beta \in [\frac{1}{(\gamma(n))^2}, 4\frac{1}{(\gamma(n))^2})$.

Note that neither h nor β are given to $\mathcal{B}$. Our goal is to construct $\mathcal{B}$ such that the acceptance probability with U and the acceptance probability with $T_{h,\beta}$ differ by a non-negligible factor. We first choose $\tilde{h}$ uniformly from the set $\{1, 2, 4, \ldots, \sqrt{N}\}$. In addition we choose δ uniformly from the range $[\frac{\sqrt{N}}{\tilde{h}}, 4\frac{\sqrt{N}}{\tilde{h}})$ and s uniformly from the range $[0, 7\frac{1}{(\gamma(n))^2})$. Then, consider the distribution $R' = C_\delta(R + Q_{\delta^2 s/N} \bmod 1)$, i.e., we first add a normal variable to R and then compress the result by a factor of δ. We take m samples $a_1, \ldots, a_m$ from $\lfloor N \cdot R' \rfloor$ and let i_0 be chosen randomly from $[m]$. We estimate $p_0(a_1, \ldots, a_m, i_0)$ and $p_u(a_1, \ldots, a_m, i_0)$ by computing many values w either according to the encryption algorithm or randomly and then calling $\mathcal{A}'$. By using a polynomial size sample, we can estimate the two probabilities up to an error of at most $\frac{1}{32n^c}$. If the two estimates differ by more than $\frac{1}{4n^c}$, $\mathcal{B}$ accepts. Otherwise, $\mathcal{B}$ rejects.

We first claim that when R is the uniform distribution, $\mathcal{B}$ rejects with high probability. The distribution $R + Q_{\delta^2 s/N} \bmod 1$ is still a uniform distribution on $[0, 1)$ and so is R' as can be easily seen from the definition of the compression. Therefore, $a_1, \ldots, a_m$ are chosen uniformly from $\{0, 1, \ldots, N-1\}$ and according to Claim A.6 if c_{m} is a large enough constant then with probability exponentially close to 1, the distribution on w obtained by encryptions of 0 is exponentially close to the uniform distribution on $\{0, 1, \ldots, N-1\}$. Therefore, since $\mathcal{A}'$ can be seen as a function on w, $|p_0(a_1, \ldots, a_m, i_0) - p_u(a_1, \ldots, a_m, i_0)|$ is also exponentially small and $\mathcal{B}$ rejects.

Now assume that R is the distribution $T_{h,\beta}$ for some fixed h and β and we claim that $\mathcal{B}$ accepts with non-negligible probability. Then, according to Claim A.3, $R + Q_{\delta^2 s/N} \bmod 1$ is $T_{h,\beta+(\delta h)^2 s/N}$. Hence, according to Claim A.5, R' is $T_{\delta h,\beta+(\delta h)^2 s/N}$. Let X denote the event that $h \le \tilde{h} < 2h$, $\delta h \in [\sqrt{N}, 2\sqrt{N})$, $\mathrm{frc}(\delta h) < \frac{1}{16m}$ and $\beta + (\delta h)^2 s/N \in [4\frac{1}{(\gamma(n))^2}, 8\frac{1}{(\gamma(n))^2})$. We now show that the probability on our choice of $\tilde{h}, \delta, s$ that X happens is at least $\frac{1}{poly(n)}$. First, with probability $\frac{1}{\log(\sqrt{N})} = \frac{2}{c_{\mathsf{N}} n^2}$, $h \le \tilde{h} < 2h$. Now, δh is uniformly distributed in $[\frac{h}{\tilde{h}}\sqrt{N}, 4\frac{h}{\tilde{h}}\sqrt{N})$. Therefore, conditioned on $h \le \tilde{h} < 2h$, the probability that $\delta h \in [\sqrt{N}, 2\sqrt{N})$ is at least $\frac{1}{3}$. Moreover, conditioned on $h \le \tilde{h} < 2h$ and $\delta h \in [\sqrt{N}, 2\sqrt{N})$, the probability that $\mathrm{frc}(\delta h) < \frac{1}{16m}$ is $\frac{1}{8m}$. For any fixed $\delta h \in [\sqrt{N}, 2\sqrt{N})$, $(\delta h)^2/N \in [1, 4)$ and therefore $\beta + (\delta h)^2 s/N$ is distributed uniformly in $[\beta, \beta + (\delta h)^2/N \cdot \frac{7}{(\gamma(n))^2})$. The length of this range is at most $4 \cdot \frac{7}{(\gamma(n))^2}$ and it always contains the range $[4\frac{1}{(\gamma(n))^2}, 8\frac{1}{(\gamma(n))^2})$ (because $\beta \in [\frac{1}{(\gamma(n))^2}, 4\frac{1}{(\gamma(n))^2})$). Therefore, the probability on the choice of s that $\beta + (\delta h)^2 s/N \in [4\frac{1}{(\gamma(n))^2}, 8\frac{1}{(\gamma(n))^2})$ is at least $\frac{4}{28} = \frac{1}{7}$. To sum up, the probability of X is at least $\frac{2}{c_{\mathsf{N}} n^2} \cdot \frac{1}{3} \cdot \frac{1}{7} \cdot \frac{1}{8m} = \frac{1}{poly(n)}$.

Notice that conditioned on X, the distribution of δh and $\beta + (\delta h)^2/N$ is the same as the distribution of h and β in the choice of the private and public keys. Therefore the probability that $(a_1, \ldots, a_m, i_0) \in Y$ is at least $Pr(X) \cdot Pr(\exists i_0, (a_1, \ldots, a_m, i_0) \in Y \mid X) \cdot \frac{1}{m} \ge Pr(X) \cdot \frac{1}{4n^c} \cdot \frac{1}{m} = \frac{1}{poly(n)}$. But when $(a_1, \ldots, a_m, i_0) \in Y$, $|p_0(a_1, \ldots, a_m, i_0) - p_u(a_1, \ldots, a_m, i_0)| \ge \frac{1}{4n^c}$ and therefore our estimates are good enough and $\mathcal{B}$ accepts. ∎

By combining the two lemmas above we get,

Theorem A.8 *For $c_{\mathsf{N}} \ge 2c_{\mathsf{h}}$ and large enough c_{m}, the public key cryptosystem described in Section 3 makes decryption errors with negligible probability and its security is based on $\sqrt{n} \cdot \gamma(n)$-uSVP.*

B Analysis of the Collision Resistant Hash Function

Claim B.1 *Let $X_1, \ldots, X_m, Y_1, \ldots, Y_m$ be mutually independent random variables. Then the statistical distance between the joint distributions satisfies $\Delta((X_1, \ldots, X_m), (Y_1, \ldots, Y_m)) \le \sum_{i=1}^m \Delta(X_i, Y_i)$.*

Proof: We consider the case $m = 2$. The claim follows for $m > 2$ by induction. According to the triangle inequality,

$$\Delta((X_1, X_2), (Y_1, Y_2)) \leq \Delta((X_1, X_2), (X_1, Y_2)) + \Delta((X_1, Y_2), (Y_1, Y_2)).$$

Since X_1 is independent of X_2 and Y_2, $\Delta((X_1, X_2), (X_1, Y_2)) = \Delta(X_2, Y_2)$ and similarly $\Delta((X_1, Y_2), (Y_1, Y_2)) = \Delta(X_1, Y_1)$. ■

Claim B.2 *Let $X_1, \ldots, X_m$ be m independent normal random variables with mean 0 and standard deviation σ. For any vector $b \in \mathbb{R}^m$, the random variable $\sum_{i=1}^m b_i X_i$ has a normal distribution with mean 0 and standard deviation $\|b\| \cdot \sigma$.*

Proof: The joint distribution $(X_1, \ldots, X_m)$ is a Gaussian distribution in $\mathbb{R}^m$ which is invariant under rotations. Hence we can equivalently consider the inner product of $(\|b\|, 0, \ldots, 0)$ and a Gaussian distribution. We complete the proof by noting that the first coordinate of the Gaussian has a normal distribution with mean 0 and standard deviation σ. ■

Definition B.3 *For any $h \in \mathbb{Z}$, $\tilde{h}, \beta \in \mathbb{R}$ and any $a \in [0, 1)$ we define the following two density functions on $[0, 1)$:*

$$S_{\tilde{h},h,\beta,a}(r) := \frac{1}{\tilde{h} \int_a^{a+\frac{1}{\tilde{h}}} T_{h,\beta}(x)dx} T_{h,\beta}(a + \frac{r}{\tilde{h}})$$

$$S'_{h,\beta,a}(r) := T_{h,\beta}(a + \frac{r}{h}) = Q_\beta(a \cdot h + r \ mod\ 1).$$

Claim B.4 *If $h \leq \tilde{h} < (1+\delta)h$ where $h \in \mathbb{Z}$, $\tilde{h} \in \mathbb{R}$, $\delta > 0$ and $\beta \leq \frac{1}{4}$ then $\Delta(S_{\tilde{h},h,\beta,a}, S'_{h,\beta,a}) \leq \frac{\tilde{c}}{\beta}\delta$.*

Proof: According to Claim D.2, $T_{h,\beta}(x) = Q_\beta(hx \bmod 1) \leq (1 + \sqrt{\beta})/\sqrt{\beta} \leq 2/\sqrt{\beta}$ for any $x \in \mathbb{R}$. Therefore,

$$\int_a^{a+\frac{1}{h}} - \int_a^{a+\frac{1}{\tilde{h}}} T_{h,\beta}(x)dx \leq \frac{2}{\sqrt{\beta}}(\frac{1}{h} - \frac{1}{\tilde{h}}) = \frac{2}{\sqrt{\beta} \cdot \tilde{h}}(\frac{\tilde{h}}{h} - 1) \leq \frac{2\delta}{\sqrt{\beta} \cdot \tilde{h}}$$

But $\int_a^{a+1/h} T_{h,\beta}(x)dx = \frac{1}{h}$ and therefore we see that

$$\frac{\tilde{h}}{h} - \tilde{h} \int_a^{a+\frac{1}{\tilde{h}}} T_{h,\beta}(x)dx \leq \frac{2\delta}{\sqrt{\beta}}.$$

Let $S''_{\tilde{h},h,\beta,a}(r) := T_{h,\beta}(a + \frac{r}{\tilde{h}})$. Then,

$$\int_0^1 \left| S_{\tilde{h},h,\beta,a}(r) - S''_{\tilde{h},h,\beta,a}(r) \right| dr = \left| 1 - \tilde{h} \int_a^{a+1/\tilde{h}} T_{h,\beta}(x)dx \right| \cdot \int_0^1 S_{\tilde{h},h,\beta,a}(r)dr =$$

$$\left| 1 - \tilde{h} \int_a^{a+1/\tilde{h}} T_{h,\beta}(x)dx \right| \leq \left| 1 - \frac{\tilde{h}}{h} \right| + \left| \frac{\tilde{h}}{h} - \tilde{h} \int_a^{a+1/\tilde{h}} T_{h,\beta}(x)dx \right| \leq (1 + \frac{2}{\sqrt{\beta}})\delta$$

Now, using the mean value theorem for any $r \in [0, 1)$,

$$\left| S'_{h,\beta,a}(r) - S''_{\tilde{h},h,\beta,a}(r) \right| \leq (\frac{1}{h} - \frac{1}{\tilde{h}}) \max_x \left| \frac{d}{dx} T_{h,\beta}(x) \right| = (\frac{1}{h} - \frac{1}{\tilde{h}}) \max_x \left| \frac{d}{dx} \sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{\beta}} e^{-\pi(\frac{h}{\sqrt{\beta}}x - \frac{1}{\sqrt{\beta}}k)^2} \right|$$

which, according to Claim D.3 using $\frac{1}{\sqrt{\beta}} \geq 2 > \frac{1}{\sqrt{2\pi}} + 1$, is at most

$$(\frac{1}{h} - \frac{1}{\tilde{h}}) \cdot \frac{\tilde{c}}{\sqrt{\beta}} \cdot \frac{h}{\sqrt{\beta}} = \frac{\tilde{c}}{\beta}(1 - \frac{h}{\tilde{h}}) \leq \frac{\tilde{c}}{\beta}\delta.$$

To sum up,

$$2\Delta(S_{\tilde{h},h,\beta,a}, S'_{h,\beta,a}) \leq \int_0^1 \left|S_{\tilde{h},h,\beta,a}(r) - S''_{\tilde{h},h,\beta,a}(r)\right| dr + \int_0^1 \left|S'_{h,\beta,a}(r) - S''_{\tilde{h},h,\beta,a}(r)\right| dr \leq (\frac{\tilde{c}}{\beta} + 1 + \frac{2}{\sqrt{\beta}})\delta \leq \frac{\tilde{c}}{\beta}\delta.$$

■

Theorem B.5 *For $c_N \geq 2c_h$ and any $c_m \geq 0$, if there exists an algorithm $\mathcal{A}$ that given a list $a_1, \ldots, a_m \in \{0, 1, \ldots, N-1\}$ finds a nonzero vector $b \in \mathbb{Z}^m$ such that $\|b\| \leq \sqrt{m}$ and $\sum_{i=1}^m b_i a_i \equiv 0(\mathrm{mod}\ N)$ with probability at least n^{-c_a} where $c_a > 0$ is some constant then there exists a solution to $\sqrt{n} \cdot \gamma(n)$-uSVP.*

Note that in particular, if $b \in \{-1, 0, 1\}^m$ then $\|b\| \leq \sqrt{m}$ and hence this theorem includes collision finding algorithms.

Proof: According to Theorem 4.1 it is enough to construct a distinguisher $\mathcal{B}$ between U and $\mathcal{T}_{\sqrt{n}\cdot\gamma(n)}$. The distinguisher $\mathcal{B}$ works by calling the routine $\mathcal{C}$ described below n times with each value $\tilde{h} = (1 + n^{-c_{\tilde{h}}})^i$, $i \in [\log_{1+n^{-c_{\tilde{h}}}} N]$. The constant $c_{\tilde{h}}$ will be specified later. If there exists an $\tilde{h}$ for which all n calls to $\mathcal{C}$ accept, $\mathcal{B}$ accepts. Otherwise, for any $\tilde{h}$ there exists one call where $\mathcal{C}$ rejects and $\mathcal{B}$ rejects.

The routine $\mathcal{C}(\tilde{h})$ samples m values $x_1, \ldots, x_m$ from the given distribution which we denote by R. It also chooses m values $y_1, \ldots, y_m$ uniformly in $[0, \frac{1}{\tilde{h}})$. Let $z_i = x_i - y_i \bmod 1$ and $a_i = \lfloor N \cdot z_i \rfloor$. We call $\mathcal{A}$ with $a_1, \ldots, a_m$. If A fails we repeat the process again (choosing x_i, y_i and calling $\mathcal{A}$). If after n^{c_a+1} calls $\mathcal{A}$ still fails, $\mathcal{C}$ accepts. Otherwise, we have a vector $b \in \mathbb{Z}^m$ such that $\|b\| \leq \sqrt{m}$ and $\sum_{i=1}^m b_i a_i \equiv 0(\mathrm{mod}\ N)$. The routine $\mathcal{C}(\tilde{h})$ accepts if $\mathrm{frc}(\sum_{i=1}^m b_i \tilde{h} y_i) < \frac{1}{4}$ and rejects otherwise.

First we show that if R is the uniform distribution then for any $\tilde{h}$, $\mathcal{C}(\tilde{h})$ accepts with probability roughly $\frac{1}{2}$. From this it will follow that the probability that n calls to $\mathcal{C}(\tilde{h})$ accept is exponentially small, i.e., $\mathcal{B}$ rejects with probability exponentially close to 1. Each number x_i is uniform in $[0, 1)$ and so is z_i. Therefore, each a_i is uniform in $\{0, 1, \ldots, N-1\}$ and according to our assumption, $\mathcal{A}$ succeeds with probability at least n^{-c_a}. The probability that n^{c_a+1} calls fail is at most $(1 - n^{-c_a})^{n^{c_a+1}} < e^{-n}$ which is exponentially small. In order to bound the probability that $\mathrm{frc}(\sum_{i=1}^m b_i \tilde{h} y_i) < \frac{1}{4}$ we use the fact that $\mathcal{A}$ is oblivious to the decomposition of the z_i's into $x_i - y_i$ and would work equally well if $z_i = x'_i - y'_i$ for some other x'_i, y'_i. Consider the following equivalent way to create the joint distribution of x_i, y_i, z_i: we first choose the z_i's uniformly in $[0, 1)$ and then choose y_i uniformly in $[0, \frac{1}{\tilde{h}})$ and choose x_i to be $z_i + y_i \bmod 1$. Hence, conditioned on any values for the z_i's, the distribution of the y_i's is uniform in $[0, \frac{1}{\tilde{h}})$ and therefore $\mathrm{frc}(\sum_{i=1}^m b_i \tilde{h} y_i)$ is distributed uniformly in $[0, \frac{1}{2})$. The probability that $\mathrm{frc}(\sum_{i=1}^m b_i \tilde{h} y_i) < \frac{1}{4}$ is therefore $\frac{1}{2}$, as required.

Now consider the case that R is $T_{h,\beta}$ where $\beta \leq \frac{4}{(\gamma(n))^2}$. We claim that when $\tilde{h}$ is the smallest such that $\tilde{h} \geq h$, $\mathcal{C}(\tilde{h})$ rejects with probability at most $\tilde{c}mn^{2c_\gamma - c_{\tilde{h}}}$. Therefore, the probability that $\mathcal{B}$ sees a rejection after n calls is at most $\tilde{c}mn^{2c_\gamma - c_{\tilde{h}}+1}$ and it therefore accepts with probability close to 1 if we choose a large enough $c_{\tilde{h}}$. Notice that such an $\tilde{h}$ satisfies $h \leq \tilde{h} < (1 + n^{-c_{\tilde{h}}})h$. As before, we create the joint distribution of x_i, y_i, z_i by first choosing z_i and then y_i. This would allow us to use the fact that $\mathcal{A}$ is oblivious to the decomposition of z_i to $x_i - y_i$. So we first choose the z_i's from their unconditional distribution and then consider the distribution of y_i conditioned on z_i given by:

$$\frac{1}{\int_{z_i}^{z_i+\frac{1}{\tilde{h}}} T_{h,\beta}(x)dx} T_{h,\beta}(z_i + r) \quad \forall r \in [0, \frac{1}{\tilde{h}}).$$

Hence the density function of the distribution of $\tilde{h} \cdot y_i$ is exactly $S_{\tilde{h},h,\beta,z_i}$. According to Claim B.4 the statistical distance between $S_{\tilde{h},h,\beta,z_i}$ and S'_{h,β,z_i} is at most $\frac{\tilde{c}}{\beta} n^{-c_{\tilde{\mathfrak{h}}}} \leq \tilde{c} n^{2c_\gamma - c_{\tilde{\mathfrak{h}}}}$. Let $\xi_1, \ldots, \xi_m$ be m random variables chosen independently according to Q_β. Notice that the distribution of the random variable $\xi_i - h \cdot z_i$ is exactly S'_{h,β,z_i}. Hence, according to Claim B.1 the statistical distance between the joint distributions $(\tilde{h} \cdot y_1, \ldots, \tilde{h} \cdot y_m)$ and $(\xi_1 - h \cdot z_1, \ldots, \xi_m - h \cdot z_m)$ is at most $\tilde{c} m \cdot n^{2c_\gamma - c_{\tilde{\mathfrak{h}}}}$. Now,

$$\sum_{i=1}^{m} b_i(\xi_i - h \cdot z_i) \bmod 1 = \sum_{i=1}^{m} b_i \xi_i - \sum_{i=1}^{m} b_i \cdot h \cdot z_i \bmod 1$$

According to Claim B.2, $\sum_{i=1}^{m} b_i \xi_i$ has a normal distribution with mean 0 and standard deviation $\|b\| \cdot \sqrt{\frac{\beta}{2\pi}} \leq \sqrt{\frac{m\beta}{2\pi}} \leq \sqrt{\frac{2m}{\pi(\gamma(n))^2}} = o(\frac{1}{\sqrt{\log n}})$. Therefore, according to Claim D.1, the probability that $\mathrm{frc}(\sum_{i=1}^{m} b_i \xi_i) > \frac{1}{8}$ is negligible. Now,

$$\mathrm{frc}(\sum_{i=1}^{m} b_i \cdot h \cdot z_i) = \mathrm{frc}(\sum_{i=1}^{m} \frac{h}{N} b_i N \cdot z_i) \leq \mathrm{frc}(\sum_{i=1}^{m} \frac{h}{N} b_i a_i) + \sum_{i=1}^{m} \frac{h}{N} \cdot b_i = \sum_{i=1}^{m} \frac{h}{N} \cdot b_i \leq m \cdot \frac{h}{N} \cdot \sqrt{m}.$$

Therefore, except with negligible probability,

$$\mathrm{frc}(\sum_{i=1}^{m} b_i(\xi_i - h \cdot z_i)) \leq \frac{1}{8} + m \cdot \frac{h}{N}\sqrt{m} < \frac{1}{4}$$

where we used the fact that $h \leq 2^{c_{\mathfrak{h}} n^2} \leq \sqrt{N}$. This implies that the probability that $\mathrm{frc}(\sum_{i=1}^{m} b_i \tilde{h} y_i) < \frac{1}{4}$ is at most $\tilde{c} m \cdot n^{2c_\gamma - c_{\tilde{\mathfrak{h}}}}$ plus some negligible amount. ■

C Quantum Computation

In this section we state a result related to a problem in quantum computation known as the dihedral hidden subgroup problem. One reason this problem is interesting is because solving it might imply a quantum solution to uSVP [16]. In [6], Ettinger and Høyer reduced the problem to the problem of finding an integer k given access to the distribution Z_k where $Pr(Z_k = z) = 2/N \cdot \cos^2(\pi k z/N)$ for $z = 0, 1, \ldots, N-1$. They presented an exponential time classical algorithm that uses only a polynomial number of samples of Z. Hence, a polynomial number of samples contains enough information to find k_0. The problem of whether there exists an *efficient* algorithm remained open.

Theorem C.1 *For any $c_1 > 0$ there exists a $c_2 > 0$ such that n^{c_2}-uSVP can be reduced to the problem of extracting k from n^{c_1} samples of Z_k.*

Proof: Omitted. ■

D Several Technical Claims

Claim D.1 *The probability that the distance of a normal variable with variance* σ^2 *from its mean is more than* t *is at most* $\sqrt{\frac{2}{\pi}} \cdot \frac{\sigma}{t} e^{-\frac{t^2}{2\sigma^2}}$.

Proof:

$$\int_t^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx \leq \int_t^\infty (1+\frac{\sigma^2}{x^2}) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = -\frac{1}{\sqrt{2\pi}\sigma} \cdot \frac{\sigma^2}{x} e^{-\frac{x^2}{2\sigma^2}} \Big|_{x=t}^{\infty} = \frac{\sigma}{\sqrt{2\pi}t} e^{-\frac{t^2}{2\sigma^2}}.$$

■

Claim D.2

$$\forall x, r \in \mathbb{R}, \quad \sum_{k \in \mathbb{Z}} e^{-\pi(kr+x)^2} \leq 1 + \frac{1}{r}$$

Proof: Let $k' \in \mathbb{Z}$ be such that $|kr + x|$ is minimized. Then,

$$\sum_{k \in \mathbb{Z}} e^{-\pi(kr+x)^2} \leq 1 + \sum_{k \in \mathbb{Z} \setminus \{k'\}} e^{-\pi(kr+x)^2} = 1 + \frac{1}{r} \sum_{k \in \mathbb{Z} \setminus \{k'\}} r \cdot e^{-\pi(kr+x)^2} \leq$$
$$1 + \frac{1}{r} \int_{-\infty}^{\infty} e^{-\pi y^2} dy = 1 + \frac{1}{r}$$

where changing the sum to an integral is possible because the sum can be seen as the area under a function that lies completely below $e^{-\pi y^2}$. ■

Claim D.3 *For any* $a, x \in \mathbb{R}$ *and any* $b > \frac{1}{\sqrt{2\pi}} + 1$, $\left| \frac{d}{dx} \sum_{k \in \mathbb{Z}} e^{-\pi(bk+ax)^2} \right| \leq \tilde{c}a$

Proof: Let z denote $a \cdot x$. Then,

$$\left| \frac{d}{dx} \sum_{k \in \mathbb{Z}} e^{-\pi(bk+ax)^2} \right| = a \left| \frac{d}{dz} \sum_{k \in \mathbb{Z}} e^{-\pi(bk+z)^2} \right| =$$
$$a \left| \sum_{k \in \mathbb{Z}} -2\pi(bk+z) e^{-\pi(bk+z)^2} \right| \leq$$
$$a \sum_{k \in \mathbb{Z}} \left| 2\pi(bk+z) e^{-\pi(bk+z)^2} \right|$$

In the following we will upper bound

$$\sum_{k \in \{0,1,\ldots\}} \left| 2\pi(bk+y) e^{-\pi(bk+y)^2} \right|$$

where $y \in [0, b]$. The upper bound for the original expression is clearly at most $2a$ times this value. The function $|2\pi y e^{-\pi y^2}|$ is increasing from 0 to $\frac{1}{\sqrt{2\pi}}$ where it attains the maximum value of $\sqrt{2\pi e}$. After that point it is monotonically decreasing. Hence,

$$\sum_{k \in \mathbb{N}} \left| 2\pi(bk+y) e^{-\pi(bk+y)^2} \right| =$$
$$\left| 2\pi y e^{-\pi y^2} \right| + \sum_{k \in \{1,2,\ldots\}} \left| 2\pi(bk+y) e^{-\pi(bk+y)^2} \right| \leq$$
$$\sqrt{2\pi e} + \int_0^\infty \left| 2\pi y e^{-\pi y^2} \right| dy = \sqrt{2\pi e} + 1$$

where changing from summation to integration is possible because $b \geq 1$ and because the function is decreasing from $\frac{1}{\sqrt{2\pi}}$ and the first y in the sum is at least $\frac{1}{\sqrt{2\pi}} + 1$. ■

Claim D.4 *For any vector $v \in L$, $\int_{\mathcal{P}(L^*)} \sum_{k\in\mathbb{Z}} e^{-\pi(\frac{k+\langle v,x\rangle}{\|v\|})^2} dx = \|v\| d(L^*)$*

Proof: First we show that the integral is invariant under basis changes of L^*, i.e., the integral does not change if we take a different fundamental parallelepiped of L^*. One can get from one basis of a lattice to another by a finite sequence of operations of the following two types: replace vector v_i by $-v_i$ and replace vector v_i by $v_i + v_j$ for some $i \neq j$. Hence, it is enough to show that the integral is invariant under these two operations. Let $f(x)$ denote $\sum_{k\in\mathbb{Z}} e^{-\pi(\frac{k+\langle v,x\rangle}{\|v\|})^2}$. Notice that for any $w \in L^*$, $f(x) = f(x+w)$ since $\langle v, w\rangle = 0$. Also, we use the following 'half' parallelepipeds:

$$\mathcal{P}_1 = \{\sum_{i=1}^{n} \alpha_i v_i^* \mid \alpha_i \in [0,1),\ \alpha_2 \geq \alpha_1\}$$

$$\mathcal{P}_2 = \{\sum_{i=1}^{n} \alpha_i v_i^* \mid \alpha_i \in [0,1),\ \alpha_2 < \alpha_1\}$$

$$\mathcal{P}_3 = \{\sum_{i=1}^{n} \alpha_i v_i^* + v_2^* \mid \alpha_i \in [0,1),\ \alpha_2 < \alpha_1\}$$

Note that $\mathcal{P}(L^*) = \mathcal{P}(v_1^*, \ldots, v_n^*)$ is equal to $\mathcal{P}_1 \cup \mathcal{P}_2$ and $\mathcal{P}(v_1^* + v_2^*, v_2^*, \ldots, v_n^*) = \mathcal{P}_1 \cup \mathcal{P}_3$. But since $\mathcal{P}_3$ is a shift of $\mathcal{P}_2$ by $v_2^* \in L^*$,

$$\int_{\mathcal{P}(L^*)} f(x)dx = \int_{\mathcal{P}_1} + \int_{\mathcal{P}_2} f(x)dx = \int_{\mathcal{P}_1} + \int_{\mathcal{P}_3} f(x)dx = \int_{\mathcal{P}(v_1^*+v_2^*, v_2^*, \ldots, v_n^*)} f(x)dx.$$

A similar argument shows that the integral is invariant under negation of basis vectors.

Now we choose $\mathcal{P}$ to be the fundamental parallelepiped created by a basis of the lattice $L^* \cap v^\perp$ and any vector w in L^* such that $\langle w, v\rangle = 1$. Then,

$$\int_{\mathcal{P}} f(x)dx = \frac{1}{\|v\|} \int_0^1 \int_{\mathcal{P}\cap\{y|\langle y,v\rangle = a\}} f(x)dx\ da =$$
$$\frac{1}{\|v\|} \int_0^1 \int_{\mathcal{P}\cap\{y|\langle y,v\rangle = a\}} \sum_{k\in\mathbb{Z}} e^{-\pi(\frac{k+a}{\|v\|})^2} dx\ da =$$
$$\frac{1}{\|v\|} \int_0^1 \|v\| d(L^*) \cdot \sum_{k\in\mathbb{Z}} e^{-\pi(\frac{k+a}{\|v\|})^2} da = d(L^*) \int_0^1 \sum_{k\in\mathbb{Z}} e^{-\pi(\frac{k+a}{\|v\|})^2} da =$$
$$d(L^*) \int_{-\infty}^{\infty} e^{-\pi(\frac{a}{\|v\|})^2} da = \|v\| d(L^*)$$

■

Properties of an LLL reduced basis

Claim D.5 *Let $B = (b_{i,j})_{1\leq i,j\leq n}$ be an $n \times n$ upper triangular matrix such that for all $i < j \leq n$, $|b_{i,j}| \leq |b_{i,i}|$. Then, the entries of $(B^T)^{-1}$ have an absolute value of at most $\frac{1}{\min_i |b_{i,i}|} 2^n$.*

Proof: First, let D denote the diagonal matrix with values $b_{i,i}$ on the diagonal. Then B can be written as MD where M is an upper triangular matrix with ones on the diagonal and all other entries have an absolute value of at most 1. Then, $(B^T)^{-1} = (D^T M^T)^{-1} = (M^T)^{-1} D^{-1}$. Therefore, it is enough to show that the entries of $L := (M^T)^{-1}$ have absolute values of at most 1. The diagonal of L is all ones and it is lower triangular. The entry $l_{i,j}$ for $i > j$ can be recursively defined by $-\sum_{j \le k < i} l_{k,j} m_{k,i}$. Therefore,

$$|l_{i,j}| = \left| \sum_{j \le k < i} l_{k,j} m_{k,i} \right| \le \sum_{j \le k < i} |l_{k,j} m_{k,i}| \le \sum_{j \le k < i} |l_{k,j}|$$

from which we get the bound $|l_{i,j}| \le 2^{i-j}$ for $i \ge j$. ∎

Claim D.6 *Let $(v_1, \ldots, v_n)$ be an LLL-reduced basis of a lattice L and let $\sum_{i=1}^n a_i v_i$ be its shortest vector. Then $|a_i| \le 2^{2n}$ for all $i \in [n]$ and $\lambda(L) \le \|v_1\| \le 2^n \lambda(L)$. Moreover, if $(v_1^*, \ldots, v_n^*)$ is the dual basis, then $\|v_i^*\| \le \frac{\sqrt{n}}{\lambda(L)} 2^{2n}$ for all $i \in [n]$.*

Proof: Let $(v_1^\dagger, \ldots, v_n^\dagger)$ denote the Gram-Schmidt orthogonalization of $(v_1, \ldots, v_n)$, i.e., $v_i^\dagger$ is the component of v_i orthogonal to the subspace spanned by $v_1, \ldots, v_{i-1}$. Clearly, $\langle v_i^\dagger, v_j \rangle = 0$ for $i > j$. Recall that in an LLL reduced basis $\|v_i^\dagger\| \le \sqrt{2} \|v_{i+1}^\dagger\|$ and for $i < j$, $\left| \langle v_i^\dagger, v_j \rangle \right| \le \frac{1}{2} \|v_i^\dagger\|^2$. In addition, recall that $\min_i \|v_i^\dagger\|$ is a lower bound on $\lambda(L)$. Then for any $i \in [n]$, $\|v_1^\dagger\| \le 2^{(i-1)/2} \|v_i^\dagger\|$ and therefore $\|v_1^\dagger\| \le 2^{(n-1)/2} \lambda(L)$. But since $v_1^\dagger = v_1$ we see that $\lambda(L) \le \|v_1\| \le 2^n \lambda(L)$. Consider the representation of $(v_1, \ldots, v_n)$ in the orthonormal basis $(v_1^\dagger / \|v_1^\dagger\|, \ldots, v_n^\dagger / \|v_n^\dagger\|)$. It is given by the columns of the matrix $B = (b_{i,j})_{1 \le i,j \le n}$ where $b_{i,j} = \langle v_j, v_i^\dagger \rangle / \|v_i^\dagger\|$. Notice that this matrix is upper triangular and that its diagonal is $b_{i,i} = \|v_i^\dagger\|$. Also note that by the properties of an LLL reduced basis, $|b_{i,j}| \le \frac{1}{2} \|v_i^\dagger\|$ for $i < j$. The shortest vector is $\sum_{i=1}^n a_i v_i = \sum_{i=1}^n (\sum_{j=i}^n a_j b_{i,j}) v_i^\dagger / \|v_i^\dagger\|$. Since its length is at most $2^n \|v_i^\dagger\|$ the absolute value of each of its coordinates is at most $2^n \|v_i^\dagger\|$. Hence, $\left| \sum_{j=i}^n a_j b_{i,j} \right| \le 2^n \|v_i^\dagger\|$ for every $i \in [n]$. By taking $i = n$ we get that $|a_n b_{n,n}| \le 2^n \|v_n^\dagger\|$ and hence $|a_n|$ is at most 2^n. We continue inductively and show that $|a_k| \le 2^{2n-k}$. Assume that the claim holds for $a_{k+1}, \ldots, a_n$. Then, $|\sum_{j=k+1}^n a_j b_{k,j}| \le \frac{1}{2} |\sum_{j=k+1}^n a_j| \|v_k^\dagger\| \le \frac{1}{2} (\sum_{j=k+1}^n 2^{2n-j}) \|v_k^\dagger\| \le \frac{1}{2} \cdot 2^{2n-k} \|v_k^\dagger\|$. By the triangle inequality, $|a_k b_{k,k}| \le |\sum_{j=k+1}^n a_j b_{k,j}| + |\sum_{j=k}^n a_j b_{k,j}| \le (\frac{1}{2} 2^{2n-k} + 2^n) \|v_k^\dagger\| \le 2^{2n-k} \|v_k^\dagger\|$ and the proof of the first part is completed.

The basis of the dual lattice is given by the columns of $(B^T)^{-1}$. Since $\min_i |b_{i,i}| \ge \frac{\lambda(L)}{2^n}$ and $|b_{i,j}| \le \frac{1}{2} |b_{i,i}|$, Claim D.5 implies that the entries of $(B^T)^{-1}$ are at most $\frac{1}{\lambda(L)} 2^{2n}$ in absolute value. Therefore, the length of each column vector is at most $\frac{\sqrt{n}}{\lambda(L)} 2^{2n}$. ∎

E Proofs

Proof of Lemma 4.2: It is convenient to have a bound on the coefficients of the shortest vector. So we assume, without loss of generality, that we are given a basis $(v_1, \ldots, v_n)$ of L which is LLL reduced. Hence, by Claim D.6, we get that the coefficients of the shortest vector satisfy $|a_i| \le 2^{2n}$ and $\frac{\|v_1\|}{2^n} \le \lambda(L) \le \|v_1\|$. These are the only properties that we need from the basis and in fact, other bases used throughout this proof will not necessarily be LLL reduced. In the following we describe a procedure $\mathcal{B}(\alpha)$ that finds the shortest vector given an estimate α which satisfies $\lambda(L) < \alpha \le 2\lambda(L)$. We apply the procedure n times with $\alpha = 2^{j-n} \cdot \|v_1\|$ for $j = 1, 2, \ldots, n+1$.

Notice that when we call $\mathcal{B}$ with the wrong value of α it can error by either outputting a non-lattice vector or a lattice vector which is longer than the shortest vector. We can easily ignore these errors by checking that the returned vector is a lattice vector and then take the shortest one. Therefore, it is sufficient to show that when α satisfies $\lambda(L) < \alpha \leq 2\lambda(L)$, $\mathcal{B}(\alpha)$ returns the shortest vector. Clearly, one can modify the $dSVP$ so that it finds whether $p \mid a_i$ for any $i \in [n]$ (and not just $i = 1$) by simply changing the order of the vectors in the basis given to the $dSVP$.

The procedure $\mathcal{B}$ is based on changes to the basis $(v_1, \ldots, v_n)$. Throughout the procedure we maintain the invariant that the lattice spanned by the current basis is a sublattice of the original lattice and that the shortest vector is unchanged. Notice that this implies that if the original lattice is an $f(n)$-unique lattice then all intermediate lattices are also $f(n)$-unique and hence all the calls to $dSVP$ are with an $f(n)$-unique lattice, as required. In addition, since the shortest vector is unchanged, the estimate α can be used whenever we call the dSVP with an intermediate lattice. The changes to the basis are meant to decrease the coefficients of the shortest vector. We let $a_1, \ldots, a_n$ denote the coefficients of the shortest vector according to the current basis. We will show that when the procedure ends all the coefficients of the shortest vector are zero except a_i for some $i \in [n]$. This implies that the shortest vector is v_i. In the following we describe a routine $\mathcal{C}$ that will later be used in $\mathcal{B}$.

The routine $\mathcal{C}(i,j)$ where $i,j \in [n]$ applies a sequence of changes to the basis. Only the vectors v_i and v_j in the basis are modified. When the routine finishes it returns the new basis and a bit. If the bit is zero then we are guaranteed that the coefficient a_i of the shortest vector in the new basis is zero. Otherwise, the bit is one and we are guaranteed that $|a_j| \leq \frac{1}{2}|a_i|$ and that a_i is nonzero. In any case, the value of $|a_i|$ does not increase by $\mathcal{C}(i,j)$.

The routine is composed of the following two steps. In the first step we replace v_i with $p \cdot v_i$ as long as the $dSVP$ says that $p \mid a_i$ and not more than $2n$ times. By multiplying v_i by p when $p \mid a_i$, we obtain a sublattice that still contains the same shortest vector. The coefficient a_i decreases by a factor of p. Since we began with $|a_i| < 2^{2n}$, if this happens $2n$ times then $a_i = 0$ and therefore in this case we return the current lattice and output a zero bit. Otherwise, we are guaranteed that in the current lattice $p \nmid a_i$.

In the second step we consider p different bases where v_i is replaced with one of $v_i - \frac{p-1}{2}v_j, \ldots, v_i - v_j, v_i, v_i + v_j, \ldots, v_i + \frac{p-1}{2}v_j$. Notice that all p bases span the same lattice. Also note that the coefficient a_j changes to $a_j + \frac{p-1}{2}a_i, \ldots, a_j + a_i, a_j, a_j - a_i, \ldots, a_j - \frac{p-1}{2}a_i$ respectively while all other coefficients remain the same. Since $p \nmid a_i$, one of the bases must satisfy that $p \mid a_j$ and we can find it by calling $dSVP_p$. We choose that basis and then multiply v_j by p. We repeat the above steps (of choosing one of the p bases and multiplying by p) $2n$ times and then output the resulting lattice with the bit one. With each step, the new $|a_j|$ becomes at most $(\frac{p-1}{2}|a_i|+|a_j|)/p = (\frac{1}{2}-\frac{1}{2p})|a_i|+\frac{|a_j|}{p}$. Hence, after $2n$ applications, the new $|a_j|$ is at most $(\frac{1}{2}-\frac{1}{2p})(1+\frac{1}{p}+\ldots+\frac{1}{p^{2n-1}})|a_i|+\frac{|a_j|}{p^{2n}} < \frac{1}{2}|a_i|+\frac{|a_j|}{p^{2n}}$ and since a_j is integer this implies $|a_j| \leq \frac{1}{2}|a_i|$. This completes the description of $\mathcal{C}$. It is easy to check that all the numbers involved have a polynomial size representation and therefore $\mathcal{C}$ runs in polynomial time.

The procedure $\mathcal{B}$ works by maintaining a set Z of possibly non-zero coefficients which is initially set to $[n]$. As long as $|Z| \geq 2$ we perform the following operations. Assume without loss of generality that $1, 2 \in Z$. We alternatively call $\mathcal{C}(1,2)$ and $\mathcal{C}(2,1)$ until the bit returned in one of the calls is zero. This indicates that one of the coefficients is zero (either a_1 or a_2 depending on which call returns the zero bit) and we remove it from the set Z. In order to show that the procedure runs in polynomial time, it is enough to show that an element is removed from Z after at most a polynomial number of steps. Notice that after each pair of calls to $\mathcal{C}$ that returned the bit one $|a_1|$ decreases by a factor of at least 4. Therefore, after at most $2n$ calls to $\mathcal{C}$, a_1 becomes zero and $\mathcal{C}(1,2)$ must return the bit zero. ■

Proof of Corollary 4.3: According to Lemma 4.2 it is enough to describe a solution to $dSVP_p$ on $p \cdot f(n)$-unique lattices. Say we are given the lattice L with the basis $(v_1, \ldots, v_n)$. By using the gap problem we can approximate $\lambda(L)$ and $\lambda(L')$ up to a factor $\sqrt{p}$ where L' is the lattice spanned by $(pv_1, v_2, \ldots, v_n)$. Notice that since $p \cdot \tau(L) \in L'$, L' is an $f(n)$-unique lattice, as required. Say that $\lambda(L) \in [d_1, \sqrt{p}d_1]$ and $\lambda(L') \in [d_2, \sqrt{p}d_2]$. If $p \mid a_1$ then both lattices contain the same shortest vector and therefore the two ranges intersect. Otherwise, it is easy to see that $\tau(L') = p \cdot \tau(L)$ and therefore the two ranges do not intersect. ■

Proof of Lemma 4.6: For any $y \in \mathbb{R}^n$,

$$|1 - \sum_{x \in L} e^{2\pi i \langle x, y \rangle} \rho(\{x\})| < \sum_{x \in L \setminus \sqrt{n}B_n} \rho(\{x\}) =$$

$$\rho(L \setminus \sqrt{n}B_n) \overset{\{1\}}{<} 2^{-\Omega(n)} \rho(L) \overset{\{2\}}{\leq} 2^{-\Omega(n)} \frac{\rho(L \cap \sqrt{n}B_n)}{1 - 2^{-\Omega(n)}} \leq 2^{-\Omega(n)}$$

where $\{1\}$ and $\{2\}$ are due to Lemma 4.4 and the last inequality holds because $\rho(L \cap \sqrt{n}B_n) = 1$. Multiplying by $d(L)$ and using Lemma 4.5 we get,

$$|d(L) - \rho(L^* + y)| < 2^{-\Omega(n)} d(L).$$

We conclude the proof by integrating over $\mathcal{P}(L^*)$,

$$\Delta(D_{L^*}, U_{L^*}) < 2^{-\Omega(n)}.$$

■

Proof of Lemma 4.7: For any $y \in \mathbb{R}^n$,

$$\left| \sum_{x \in L} e^{2\pi i \langle x, y \rangle} \rho(\{x\}) - \sum_{k \in \mathbb{Z}} e^{2\pi i k \langle u, y \rangle} \rho(\{ku\}) \right| <$$

$$\sum_{x \in L \setminus \sqrt{n}B_n} \rho(\{x\}) = \rho(L \setminus \sqrt{n}B_n) \overset{\{1\}}{<} 2^{-\Omega(n)} \rho(L) \overset{\{2\}}{\leq} 2^{-\Omega(n)} \frac{1}{1 - 2^{-\Omega(n)}} \rho(\{ku | k \in \mathbb{Z}\}) \leq$$

$$2^{-\Omega(n)} \rho(\{ku | k \in \mathbb{Z}\}) \overset{\{3\}}{\leq} 2^{-\Omega(n)} (1 + \frac{1}{\|u\|})$$

where $\{1\}$ and $\{2\}$ are due to Lemma 4.4 and $\{3\}$ is due to Claim D.2 with $x = 0$. By multiplying by $d(L)$ we get

$$|\rho(L^* + y) - d(L) \sum_{k \in \mathbb{Z}} e^{2\pi i k \langle u, y \rangle} \rho(\{ku\})| < 2^{-\Omega(n)} (1 + \frac{1}{\|u\|}) \cdot d(L). \tag{2}$$

Consider the one dimensional lattice M spanned by the number $\|u\|$. Clearly, the lattice M^* is spanned by the number $\frac{1}{\|u\|}$. According to Lemma 4.5 for any $a \in \mathbb{R}$,

$$\rho(M^* + a) = d(M) \sum_{b \in M} e^{2\pi i a b} \rho(\{b\}) = \|u\| \sum_{k \in \mathbb{Z}} e^{2\pi i k a \|u\|} \rho(\{ku\}).$$

Therefore,

$$d(L)\sum_{k\in\mathbb{Z}} e^{2\pi ik\langle u,y\rangle}\rho(\{ku\}) = \frac{d(L)}{\|u\|}\rho(M^* + \frac{\langle u,y\rangle}{\|u\|}) = T_{L^*,u}(y)$$

We conclude the proof by integrating (2) over $\mathcal{P}(L^*)$:

$$\Delta(D_{L^*}, T_{L^*,u}) < 2^{-\Omega(n)}(1 + \frac{1}{\|u\|}).$$

■

Proof of Lemma 4.8: We are given a basis $(v_1, \ldots, v_n)$ of a $g(n)$-unique lattice L and a number α such that $\lambda(L) < \alpha \le 2\lambda(L)$. Let L' be the lattice L scaled by a factor $\frac{2\sqrt{n}}{\alpha\cdot g(n)}$, i.e., the lattice spanned by the basis $(v'_1, \ldots, v'_n) := \frac{2\sqrt{n}}{\alpha\cdot g(n)}(v_1, \ldots, v_n)$. Notice that in L' the shortest vector $\tau(L') = \sum_{i=1}^n a_i v'_i$ is of length in $[\frac{\sqrt{n}}{g(n)}, \frac{2\sqrt{n}}{g(n)})$ and any vector not parallel to $\tau(L')$ is of length at least $g(n)\cdot\frac{\sqrt{n}}{g(n)} = \sqrt{n}$. Now, let M be the lattice spanned by the basis $(p(n)v'_1, v'_2, \ldots, v'_n)$. If $p(n) \mid a_1$ then $\tau(M) = \tau(L')$ and therefore its length is in $[\frac{\sqrt{n}}{g(n)}, \frac{2\sqrt{n}}{g(n)})$. Also, since $M \subseteq L'$, any vector in M not parallel to $\tau(M)$ is of length at least $\sqrt{n}$. If $p(n) \nmid a_1$ then the shortest multiple of $\tau(L')$ that is contained in M is $p(n)\cdot\tau(L')$ whose length is at least $p(n)\cdot\frac{\sqrt{n}}{g(n)} > \sqrt{n}$. Hence, in this case all non-zero vectors are of length at least $\sqrt{n}$.

Clearly, we can take an LLL reduced basis of the lattice M without changing the properties of the lattice described above. Now consider the distribution D_{M^*}. We can efficiently sample from it by sampling a Gaussian centered around the origin with standard deviation $\frac{1}{\sqrt{2\pi}}$ and reducing it modulo $\mathcal{P}(M^*)$. According to Lemma 4.7, if $p(n) \mid a_1$ then the distribution is exponentially close to $T_{M^*,\tau(M)}$. On the other hand, if $p(n) \nmid a_1$, Lemma 4.6 says that the distribution is exponentially close to the uniform distribution U_{M^*}. Therefore, we can decide with non-negligible probability if $p(n) \mid a_1$ by calling an algorithm that distinguishes between $T_{M^*,\tau(M)}$ and U_{M^*}. The error probability can be made exponentially small by calling the algorithm a polynomial number of times and taking the majority. ■