Electronic Colloquium on Computational Complexity - Reports Series 1996 - available via:

Revision 01 of ECCC TR96-065

FTP: ftp.eccc.uni-trier.de:/pub/eccc/
WWW: http://www.eccc.uni-trier.de/eccc/
Email: ftpmail@ftp.eccc.uni-trier.de with subject 'help eccc'

A Public-Key Cryptosystem with Worst-Case/Average-Case Equivalence

(Extended Abstract)

Miklos Ajtai Cynthia Dwork

May 2, 1997

Abstract

We present a probabilistic public key cryptosystem which is secure unless the worst case of the following lattice problem can be solved in polynomial time: "Find the shortest nonzero vector in an n dimensional lattice L where the shortest vector v is unique in the sense that any other vector whose length is at most $n^c\|v\|$ is parallel to v."

1 Introduction

The *unique shortest vector problem (u-SVP)* is to find the shortest nonzero vector in an n dimensional lattice L where the shortest vector v is unique in the sense that any other vector whose length is at most $n^c\|v\|$ is parallel to v. We present a public key cryptosystem generator with the property that if a random instance of the cryptosystem can be broken, that is, if for a random instance the probability that an encryption of a zero can be distinguished from an encryption of a one (without the private key) in polynomial time is at least $\frac{1}{2} + n^{-c_1}$ for some absoloute constant $c_1 > 0$, then the worst-case unique shortest vector problem has probabilistic polynomial time solution[1]. To our knowledge this is the first public key cryptosystem with the property that to break a random instance is as hard as to solve the worst-case instance of the problem on which the system is based.

Our approach also yields a conceptually simple and extremely natural pseudo-random generator.

Outline of the Construction Very roughly speaking, an instance of the cryptosystem is a collection of *hidden hyperplanes*, which form the private key, together with a method of generating a point guaranteed to be near one of the hyperplanes in the collection, which forms

[1]The unique shortest vector problem is one of the three problems listed in [2]. There, a random method is given to generate hard instances of a particular lattice problem so that if it has a polynomial time solution then all of the three worst-case problems (including the unique-shortest vector problem) has a solution.

the public key. The public key is chosen so as not to reveal the collection of hyperplanes – indeed, any ability, given the public key, to discover the collection implies the ability to solve the worst-case unique shortest vector problem. Encryption is bit-by-bit: zero is encrypted by using the public key to find a random vector $v \in \mathbb{R}^n$ near one of the hyperplanes – the ciphertext is v; one is encrypted by choosing a random vector u uniformly from $\mathbb{R}^n$ – the ciphertext is simply u. Decryption of a ciphertext x is performed using the private key to determine the distance of x to the nearest hidden hyperplane. If this distance is sufficiently small, then x is decrypted as zero; otherwise x is decrypted as one. There is a small (but polynomial) probability of an error in decryption: an encryption of one may be decrypted as zero.

We present three separate cryptosystems, the last of which is the system just described. The first has the most compact public key; its correctness depends on the hardness of random instances of a certain subset of instances of the unique shortest vector problem. The second has a less compact public key, but its correctness depends only on the hardness of random instances of the unique shortest vector problem (that is, we are no longer restricted to a subset). The third has the least compact public key. However, its correctness relies only on the hardness of the worst-case unique shortest vector problem.

In all three constructions the hyperplanes are obtained by regarding the unique shortest vector u in lattice Λ as a linear functional inducing the collection of hyperplanes $H_i = \{v \mid u \cdot v = i\}$ for each $i \in \mathbb{Z}$. The private key is a basis for H_0. Every point in $L = \Lambda^*$ (the dual of Λ) is on one of the H_i. In the first two constructions the public key is a random basis for L, together with an additional parameter R. In these schemes, a point near a hyperplane (an encryption of zero) is obtained by choosing a random point in L and perturbing it slightly (using R). In the third construction the public key is a collection of random points, themselves near the H_i, together with the parameter R. A point near a hyperplane is obtained by perturbing the sum of a random subset of the published points.

In this Extended Abstract all proofs have been omitted for lack of space. Full proofs appear in the Appendix.

2 Definitions

The fundamental concepts concerning lattices and public-key crytposystems can be found in [6, 11, 12, 13, 9, 17].

A *lattice* in $\mathbb{R}^n$ is a set of the form

$$L = L(b_1, \ldots, b_n) = \{\sum_{i=1}^{n} \lambda_i b_i \mid \lambda_i \in \mathbb{Z}, i = 1, \ldots, n\},$$

where $b_1, \ldots, b_n$ is a basis of $\mathbb{R}^n$. We say that $(b_1, \ldots, b_n)$ is a *basis* of L. The *length* of a vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, denoted $\|x\|$ is $(x_1^2 + \ldots + x_n^2)$. The length of the basis $(b_1, \ldots, b_n)$ is the length of the longest basis vector, $\max_{i=1}^{n} \|b_i\|$. The *determinant* of L, denoted $\det(L)$, is the absolute value of the determinant of the parallelepiped with sides $b_1, \ldots, b_n$, where $b_1, \ldots, b_n$ is *any* basis for the lattice: $\det(L) = |\det(b_1, \ldots, b_n)|$.

The *dual* lattice of L, denoted L^*, is defined as

$$L^* = \{x \in \mathbb{R}^n \mid x^T y \in \mathbb{Z} \text{ for all } y \in L\}.$$

If $(b_1, \ldots, b_n)$ is a basis of L then $(c_1, \ldots, c_n)$ is a basis for L^*, where

$$c_i^T b_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

If $a_1, \ldots, a_n \in \mathbb{R}^n$ are linearly independent vectors, then $\mathcal{P}^-(a_1, \ldots, a_n)$ denotes the half-closed parallelepiped

$$\{\sum_{i=1}^{n} \gamma_i a_i \mid 0 \leq \gamma_i < 1, i = 1, \ldots, n\}.$$

Let P_1 and P_2 denote two probability distributions and let Ω be a σ-field. The *distance* between P_1 and P_2 is

$$\sup\{|P_1(A) - P_2(A)| + |P_1(B) - P_2(B)| \text{ s.t. } A, B \in \Omega,\ A \cap B = \emptyset\}.$$

The n *dimensional ball of radius* R is the set of vectors $x \in \mathbb{R}^n$ such that $\|x\| \leq R$.

3 (d, M)-Lattices

Assume n is a positive integer, $M > 0$, $d > 0$ are real numbers, and $L \subseteq \mathbb{Z}^n$ is a lattice which has an $n-1$ dimensional sublattice L' with the following properties:

1. L' has a basis of length at most M;
2. if H is the $n-1$ dimensional subspace of $\mathbb{R}^n$ containing L' and $H' \neq H$ is a coset of H intersecting L, then the distance of H and H' is at least d.

Then we say that L is a (d, M)-lattice. If $d > M$, then L' is unique. In this case L' will be denoted by $L^{(d,M)}$. The minimum distance between H and a coset of H intersecting L will be denoted d_L.

Let $c > 5$ be a real number, and let $\mathcal{L}$ be a distribution on the set of (d, M) lattices for which $d > n^c M$ and $d \leq d_L \leq 2d$. The *hidden hyperplane* assumption for $\mathcal{L}$ says that, given a basis for a random (d, M) lattice $L \in_R \mathcal{L}$, it is computationally infeasible to compute $L^{(d,M)}$.

The hidden hyperplane assumption is related to the unique shortest vector problem as follows. If Λ is a lattice with an n^c-unique shortest vector u, then $L = \Lambda^*$ is a (d, M) lattice for some $d \geq n^c M$ and $d_L = \|u\|^{-1}$. Let $H = H_0$ be the $n-1$ dimensional subspace of $\mathbb{R}^n$ containing $L^{(d,M)}$, and in general let $H_i = \{v \mid u \cdot v = i\}$. Then u is orthogonal to H (because the inner product of u with any vector in H is 0), and the distance between adjacent H_i is $\|u\|^{-1}$. Thus, knowing H reveals the direction of u, and, by computing the gcd of the distances to H of random points in L, d_L can be computed in probabilistic polynomial time, yielding $\|u\|$ (and hence, u).

In all three cryptosystems the value one is encrypted by choosing a random point in a particular region in $\mathbb{R}^n$ (the exact depends on the scheme).

In the first cryptosystem, the public key is a random basis of a (d, M) lattice where the length of the random basis is greater than d_L by only a polynomial (in n) factor. Using this constraint on the length of the basis in the public key, we prove that the ability to distinguish encryptions of zero from encryptions of one yields the ability to find H (and hence, u).

In the second cryptosystem we remove the constraint on the length of the published basis for L. We show that the ability to distinguish encryptions of zero from encryptions of one yields the ability to construct $n-1$ mutually orthogonal long vectors very close to H, from which it is possible to find H exactly (and hence, u).

In the third cryptosystem no lattice is presented; rather, the public key is a set of random points near the hyperplanes induced by a random vector u in the n-dimensional unit ball. The sum of a random subset of these points is itself close to a hyperplane; an encryption of zero is a perturbation on such a random subset sum, reduced modulo a certain parallelepiped determined by the public key. We introduce additional machinery to prove that every instance of the n^c-unique shortest vector problem can with overwhelming probability be efficiently transformed into a random instance of the third cryptosystem, for which the ability to distinguish encryptions of zero from encryptions of one yields the unique shortest vector. This proof is built on the results that we proved about the first and second systems.

4 First and Second Cryptosystems

The Key Pair Generation Procedure

1. Generate a random $n-1$ dimensional lattice L' which has a basis $(b_1, \ldots, b_{n-1})$ such that $\|b_i\| \leq M$; for example, we can use the random class given in [2]. Let H be the $n-1$ dimensional subspace containing L'.

2. Choose $d \geq n^c M$.

3. Choose from a large cube a random vector b_n of distance $d \leq d_L \leq 2d$ from H.

4. The private key is any basis for $L^{(d,M)}$ (equivalently, for H).

5. Construct a random basis B' for $L = L(B)$. The public key is (B', M).

The Encryption and Decryption Procedures

To encrypt zero, choose a random lattice point v in the cube KU^n, where U^n is the n dimensional unit cube and $K \geq 2^n d$. For $R \in \mathbb{R}$ and $m \in \mathbb{Z}$, let the *perturbation* $\text{pert}(R, m)$ be the random variable whose value is the sum of m vectors taken independently and with uniform distribution from the n dimensional ball with radius R around 0. For $m = c_0 n$, $c_0 \geq 4$, and $R = n^3 M$, choose a value w of $\mathrm{pert}(R, m)$. The ciphertext is $v + w$. To encrypt one, choose a random (probably non-lattice) point in KU^n; this point is the ciphertext.

Let u_H be a unit vector orthogonal to the subspace H, and let d_L be the distance between consecutive hyperplanes. To decrypt the ciphertext z, the receiver computes the fractional part of $(u_H \cdot z)/d_L$. If it is within $\frac{mR}{d_L}$ of 0 or 1 then z is decrypted as 0, and as 1 otherwise.

5 Reduction for the First System

In this section we outline the proof that that if we constrain the distribution $\mathcal{L}$ so that each (d, M) lattice $L \in \mathcal{L}$ can be presented by a basis whose length exceeds d_L by at most a polynomial (in n) factor, then the ability to distinguish encryptions of zero from encryptions of one yields the ability to solve the hidden hyperplane problem. This implies that the only way to break the cryptosystem is to find the private key.

Following [2], we assume there is a procedure, which, given a basis Y for L, samples lattice points within a cube whose side has length at least $n^c \|Y\|$ with distribution exponentially close to uniform. We also frequently need to choose a vector t uniformly from $S^n(R)$. This is done inductively, one coordinate at a time, beginning with the nth coordinate.

5.1 Indistinguishability of Distributions

Let L be a lattice and let $K > 0, R > 0$ be real numbers. The random variable $\xi_{L,K,R}$ is defined in the following way: we choose a point x uniformly at random from $KU^{(n)} \cap L$, where $U^{(n)}$ is the unit cube in $\mathbb{R}^n$ and we choose a value w of $\mathrm{pert}(R, m)$, where $m = c_0 n$ for some $c_0 \geq 4$. The value of $\xi_{L,K,R}$ is $x + w$.

η_K will be a random variable whose values are taken with uniform distribution on $KU^{(n)}$. Choose $\delta \in_R \{0, 1\}$. $\nu_{L,K,R}$ is defined in the following way. We randomize $\delta, \xi_{L,K,R}$ and η_K independently. If $\delta = 0$, then $\nu_{L,K,R} = \eta_K$, if $\delta = 1$ then $\nu_{L,K,R} = \xi_{L,K,R}$.

Suppose that the real number $c > 5$ and the positive integers $n, d, M, K, R,\ d > n^c M$ are given, and $\mathcal{L}$ is a distribution on (d, M)-lattices in $\mathbb{Z}^n$. We say that a probabilistic algorithm $\mathcal{A}$ finds $L^{(d,M)}$ on $\mathcal{L}$ with a probability p, if given as input a description of $\mathcal{L}$ (including d and M) and $L \in \mathcal{L}$, $\mathcal{A}$ outputs $L^{(d,M)}$ with probability p, where the probability is taken both for the randomization of L and for the randomization in $\mathcal{A}$. Sometimes we will allow $\mathcal{A}$ to use an oracle. In this case each use of the oracle will be counted as one time unit in the definition of the time used by $\mathcal{A}$.

We assume a model in which for some constants e_0 and e_1, a $2^{-n^{e_0}}$ approximation to a real can be obtained in time $\Theta(n^{e_1})$. Suppose that the real number $c > 5$ and the positive integers $n, d, M, K, R,\ d > n^c M$ are given, and $\mathcal{L}$ is a distribution on (d, M)-lattices in $\mathbb{Z}^n$. We say that the probabilistic algorithm $\mathcal{A}$ *distinguishes* $\xi_{L,K,R}$ *and* η_K *on* $\mathcal{L}$ *with a probability* p if given a description of $\mathcal{L}$, $L \in_R \mathcal{L}$, and a random value of $\nu_{L,K,R}$ as an input (together with n, M, d, K, R), $\mathcal{A}$ outputs a 0, 1 value w so that $P(w = \delta) = p$. Note that in polynomial time $\mathcal{A}$ sees only polynomial (in n) bits of its input.

Theorem 5.1 *There exist $c, c_4, c_5, c_6 > 0$ so that for all $c_1 > 0$, $c_2 > 0$ there exists $c_3 > 0$ and a probabilistic algorithm $\mathcal{B}$ (using an oracle) so that if n, d, M, K, R are positive integers satisfying the inequalities,*
(1) $\log d + \log M + \log K + \log R < n^{c_1}$,
(2) $d > n^{c_6} M$,
(3) $R > n^c M$,
(4) $2^{c_5 n} d > K > 2^{c_4 n} d$,
and $\mathcal{L}$ is a distribution on the set of (d, M) lattices in $\mathbb{Z}^n$ presented by vectors of length at most $n^{c_1} d_L$ and for which $d_L > n^5 M$ and $d \le d_L \le 2d$, and $\mathcal{A}$ is a probabilistic algorithm which distinguishes $\xi_{L,K,R}$ and η_K on $\mathcal{L}$ with probability at least $\frac{1}{2} + n^{-c_2}$, then $\mathcal{B}$, using $\mathcal{A}$ as an oracle, finds $L^{(d,M)}$ on $\mathcal{L}$ with a probability at least $1 - 2^{-n}$, in time n^{c_3}.

Let $L \in_R \mathcal{L}$ be presented by $(b_1, \ldots, b_n)$ such that $\max_{1 \le i \le n} \|b_i\| \le n^{c_1} d_L$. Strictly speaking, as described above, we must charge time $\theta(n^{e_1})$ for $\mathcal{B}$ to access a $2^{e_0 n}$ approximation to a real input. For simplicity, we describe $\mathcal{B}$ as if it and $\mathcal{A}$ could access any real in a single step (this issue is addressed in the Appendix).

Algorithm $\mathcal{B}$ works as follows. Let $K' = n^c d_L$. Choose a polynomial (in n) random lattice points $p_1, \ldots, p_{m'} \in K'U^n$. (This is where we use the assumption that L is presented by a basis of length at most polynomial in n larger than d.) For $1 \le i < j \le m'$, let $a_{ij} = p_i - p_j$.

Note that $K'U^n$ is intersected by at most n^c cosets H' of H intersecting L, where H is the $n-1$ dimensional subspace containing $L^{(d,M)}$. Let H' be a coset of H whose intersection with $K'U^n \cap L$ is maximal. The number of differences a_{ij} such that p_i and p_j are both in H' is at least $(\frac{1}{n^c})^2(\frac{m'(m'-1)}{2})$, so a polynomial fraction of the a_{ij} are in H. The key idea, described below, is to use $\mathcal{A}$ to determine which of the differences a_{ij} are in H. By doing so, if m' is sufficiently large, then, by arguments appearing in [2], $\mathcal{B}$ will find a basis for H among the a_{ij}.

Testing for Containment in H

Let $L^{(d,M)} = L(b_1, \ldots, b_{n-1})$, where $\max_{1 \le i \le n-1} \|b_i\| \le M$, and let $\mathcal{P}' = \mathcal{P}^-(b_1, \ldots, b_{n-1})$. Let w be a value of pert(R, m). We prove that if $\epsilon > 0$ is arbitrary (it may depend on n) and σ is a strip of width ϵ, not too far from the hyperplane H, then the distribution of the projection of w reduced modulo $(b_1, \ldots, b_{n-1})$ into the lattice parallelepiped $\mathcal{P}'$ is almost uniform on the hyperlane, even with the condition that w is in σ.

Each $v \in \{a_{ij} \mid 1 \le i < j \le m'\}$ induces a probability distribution as follows. Let u be a random variable with uniform distribution on $KU^n \cap L$, and let α be a random variable distributed uniformly in $[0, 1]$. Define the random variable $\delta_v = u + \alpha v + w$. It follows from the uniformity of the projection of w onto H (modulo $\mathcal{P}'$) that the distributions obtained by projecting δ_v and $\xi_{L,K,R}$ onto H are almost uniformly distributed on the projection of KU^n on H, independent of whether or not $v \in H$. Moreover, this is true even if we restrict the distributions to the case in which w lies in a strip σ not too far from H.

If $v \in H$, then $u + \alpha v \in H'$, where H' is the coset of H containing $u \in L$. In this case, if v is not too long, then $u + \alpha v + w$ has essentially the same distribution as $\xi_{L,K,R}$: each depends only on the distance from H of its respective copy of pert(R, m). If $v \notin H$, then since with all but exponentially small probability u and v are not in the same coset of H, the signed distance of $u + \alpha v$ to the nearest coset of H is uniformly distributed in $(-\frac{d_L}{2}, \frac{d_L}{2}]$; so if v is not too long then δ_v has essentially the same distribution as η_K. Thus, the assumed ability of $\mathcal{A}$ to distinguish $\xi_{L,K,R}$ from η_K reveals whether or not $v \in H$.

Remark. The indistiguishability of $\xi_{L,K,R}$ from η_K yields a pseudo-random number generator. Each of the three cryptosystems yields a generator in this way.

6 Extension to General Lattices

As before, we wish to show that if there is a probabilistic polynomial time machine that distinguishes $\xi_{L,K,R}$ from η_K, then, using the distinguisher, H can be found in probabilistic polynomial time. However, if the lattice L is presented by a basis of length greater than $n^c d$ for some $c > 0$ then the previous reduction fails: we can no longer sample lattice points inside a small cube. We get around this problem by using the distinguisher to help us find random short vectors very close to H, and then then "growing" these into long vectors, still quite close to H. The growing takes place in stages; we use the distinguisher at every stage to recognize when a vector close to H has been found.

The long vectors are then used to find an approximation to H. If the approximation is sufficiently good then the unit vector orthogonal to the approximation will be very close to the unit vector u_H orthogonal to H. If the two unit vectors are sufficiently close then u_H can be found by rounding the unit vector orthogonal to the approximation.

Growing Long Vectors

Let σ denote the set of all points in $\mathbb{R}^n$ of distance at most d from H. Each iteration has a *starting point* s which for the first iteration is the origin, and in general will always be within distance $2d$ of H. Let $S(2\sqrt{n}d, s)$ be a ball of radius $2\sqrt{n}d$ around the starting point s. The goal is to find a point v in $\sigma \cap S(2\sqrt{n}d, s)$ that is farther from the origin than s and still inside σ. Then $2v$ becomes the new starting point, and the process continues. Occasionally the procedure may err; this is eventually detected and the computation is backed up to an earlier starting point and repeated with different random choices.

In Section 5.1 we used the distinguisher to test points $v \in L$ to see if they are outside of H. Specifically, v was tested by sampling the distribution δ_v and testing $\mathcal{A}$ on the samples. We will use the same test here, this time to distinguish points near H from points outside of σ. Specifically, we have a way of choosing random points v within distance $2d$ of H and testing them such that: (1) if $v \notin \sigma$ then with high probability this is detected; (2) if v is "very close" to H then with high probability v is recognized as being in σ; and (3) the probability that we find a $v \in \sigma$ that is not falsely detected as being outside of σ is polynomial in n^{-1}.

Our goal is to construct an approximation $\hat{H}$ to H by finding $n-1$ mutually orthogonal long lattice vectors $v_1, \ldots, v_{n-1}$, say, of length at least ℓ for a suitably chosen ℓ, all at distance less than d of H. Once we have found $v_1, \ldots v_{i-1}$, we search for v_i in the $n-i+1$ dimensional subspace V^{n-i+1} of $\mathbb{R}^n$ orthogonal to $v_1, \ldots, v_{i-1}$, such that v_i is close to $H \cap V^{n-i+1}$.

We now describe the general step of searching for the next starting point in the construction of a vector of length ℓ within distance d_L of H.

Finding the Next Starting Point
Choose a random $v' \in S^n(2\sqrt{n}d)$ such that $||s+v'|| > ||s||$. Test if $s+v'$ is outside of σ, by testing each of $s+v', s+\frac{n^{c_1}-1}{n^{c_1}}v', s+\frac{n^{c_1}-2}{n^{c_1}}v', \ldots s+\frac{1}{n^{c_1}}v'$ to see if any is near a coset $H_1 \neq H$ in L. (For any vector u this test is accomplished by sampling from δ_u.) If any multiple of v' tests positive, then v' is discarded and the procedure is repeated for a new random v'. If no test is positive and $\|s+v'\| \geq \ell$, then we set $v = s+v'$. If no test is positive but $\|s+v'\| < \ell$, then we set the next starting point to $2(s+v')$.

Theorem 6.1 *There exist $c, c_4, c_5, c_6 > 0$ so that for all $c_1 > 0$, $c_2 > 0$ there exists $c_3 > 0$ and a probabilistic algorithm $\mathcal{B}$ (using an oracle) so that if n, d, M, K, R are positive integers satisfying the inequalities,*
(1) $\log d + \log M + \log K + \log R < n^{c_1}$
(2) $n^{c_6}M > d > n^{c_4}M$,
(3) $R > n^c M$,
(4) $K > 2^{c_5 n} d$,
and $\mathcal{L}$ is a distribution on (d, M)-lattices in $\mathbb{Z}^n$ presented by vectors in a cube of size $2^{n^{c_1}} d$, and $\mathcal{A}$ is a probabilistic polynomial time algorithm which distinguishes $\xi_{L,K,R}$ and η_K on $\mathcal{L}$ with a probability of at least $\frac{1}{2} + n^{-c_2}$, then, $\mathcal{B}$, using $\mathcal{A}$ as an oracle, finds $L^{(d,M)}$ on $\mathcal{L}$ with a probability of at least $1 - 2^{-n}$, in time n^{c_3}.

Remark. Since M is just an upper bound on the length of $L^{(d,M)}$ the requirement that $n^{c_6}M > d$ does not restrict $\mathcal{L}$.

7 The Main Theorem: Worst-Case/Average-Case Equivalence

In this section we will use three constants, $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$. We assume that $\mathcal{D}_1 = 3, \mathcal{D}_2 = 8, \mathcal{D}_3 = 3$. In a similar way $\mathcal{K}(n)$ will denote function $2^{n \log n}$. We have made no attempt to choose these constants and the function in an optimal way in any sense.

As mentioned earlier, in the third cryptosystem the public key involves no lattice. Suppose that $u \in \mathbb{R}^n$, $0 < \|u\| \leq 1$, $R > 0$ and m is a positive integer. Let $\mathcal{Q}$ be the n-dimensional cube $\mathcal{K}U^{(n)}$. We define the random variable $\mathcal{H}'(u, R, m)$ in the following way: First let X be the set of all $x \in \mathcal{Q}$ so that $x \cdot u$ is an integer. X consists of subsets of a finite number of $n-1$-dimensional hyperplanes, so the $n-1$ dimensional volume defined on these hyperplanes induces a probability measure on X. We take a random point y on X. Independently we also take a value z of $\mathrm{pert}(R, m)$. The value of $\mathcal{H}'(u, R, m)$ is $y + z$. Let $\mathcal{H} = \mathrm{round}_{2^{-n}}(\mathcal{H}')$, where for $y \in \mathbb{R}$ and $\alpha > 0$, $\mathrm{round}_\alpha(y) = i\alpha$, where i is the largest integer with $i\alpha \leq y$ and if $x = \langle x_1, ..., x_n \rangle \in \mathbb{R}^n$ then $\mathrm{round}_\alpha(x) = \langle \mathrm{round}_\alpha(x_1), ..., \mathrm{round}_\alpha(x_n) \rangle$.

In the third system the private key is a random vector u chosen with uniform distribution on the set $\{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$. The public key is a set of m independent values $v_1, ..., v_m$ of the random variable $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$, where $m = n^{\mathcal{D}_3}$, so by definition, the values v_i in the public key are small perturbations of points in the hyperplanes induced by u.

For the encryption of a message, the sender will need the smallest integer i_0 so that $\mathrm{width}(v_{i_0+1}, \ldots, v_{i_0+n})$ is at least $n^{-2}\mathcal{K}$, where if $a_1, ..., a_n \in \mathbb{R}^n$, then $\mathrm{width}(a_1, ..., a_n)$ is the width of the parallelepiped defined by the vectors $a_1, ..., a_n$ (that is, the maximum of the distances between the point a_i and the subspace generated by $\{a_j | j \neq i\}$, for $i = 1, ..., n$). We prove that, with a probability exponentially close to 1, $i_0 < n^2$. Since the value of i_0 does not depend on the message, we may consider i_0 to be part of the public key. Let $\mathcal{P} = \mathcal{P}(v_{i_0+1}, ..., v_{i_0+n})$.

An encryption of zero is obtained by computing the vector $x = \sum_{j=1}^{m} \delta_i v_i$, where each $\delta_i \in_R \{0,1\}$, and reducing x modulo $v_{i_0+1}, ..., v_{i_0+n}$ into $\mathcal{P}^-$, that is, finding the unique vector x' in $\mathcal{P}^-(v_{i_0+1}, ..., v_{i_0+n})$ so that $x - x'$ is an integer linear combination of the vectors $v_{i_0+1}, ..., v_{i_0+n}$. x' is the ciphertext. Let the random variable $\mathcal{S}_{v_1,...,v_m}$ be a random encryption of zero, as just described. An encryption of one is obtained by choosing a random point in $\mathcal{P}^- \cap 2^{-n}\mathbb{Z}^n$, where $2^{-n}\mathbb{Z}^n$ is the set of all vectors of the form $2^{-n}b$, $b \in \mathbb{Z}^n$. Let the random variable $\mathcal{E}_{v_1,...,v_m}$ be a random encryption of one, as just described. In light of the results for the first two cryptosystems, the intuition for indistinguishability of the two distributions is that encryptions of zero are themselves relatively small perturbations of points on the hyperplanes induced by u, while encryptions of one are just random points in space. For this system, however, we obtain the following worst-case/average-case hardness result.

Theorem 7.1 *For all $c_1, c_2, c_3, c_4 > 0$ there exists a c_5 and a probabilistic algorithm $\mathcal{B}$ (using an oracle) so that for all sufficiently large n, condition (1) implies condition (2), where*

(1) $\mathcal{A}$ is a probabilistic circuit of size n^{c_1} so that if $u, v_1, ..., v_m$ are picked at random as described in the protocol for generating the public and private keys, then with a probability of at least n^{-c_2} the following holds:

$\mathcal{A}$ distinguishes the random variables $\mathcal{S}_{v_1,...,v_m}$ and $\mathcal{E}_{v_1,...,v_m}$, given $v_1, ..., v_m$, with probability at least $\frac{1}{2} + n^{-c_3}$.

(2) $\mathcal{B}$, using $\mathcal{A}$ as an oracle, can solve any instance of size at most n^{c_4} of the $n^{\mathcal{D}_2}$-unique shortest vector problem in time n^{c_5} and with a probability at least $1 - 2^{-n}$.

The proof of the theorem is in two parts. In the first part, assume that there exists a probabilistic polynomial time machine $\mathcal{A}$ that, given the public key, followed by a block of t random encryptions of the bit b, followed by a block of t random encryptions of $1 - b$, produces b with probability polynomially better than $1/2$ for a polynomial fraction of the instances of the cryptosystem (an instance is a (public key, private key) pair generated by the cryptosystem generator). Let $\mathcal{U}'$ be a random variable which takes its values with uniform distribution on the n-dimensional cube $\mathcal{K}U^{(n)}$ and let $\mathcal{U} = \mathrm{round}_{2^{-n}}(\mathcal{U}')$. We show that the existence of $\mathcal{A}$ implies the existence of a probabilistic polynomial time machine $\mathcal{C}$ that, on input mt values of a random variable ξ, using $\mathcal{A}$ as an oracle, determines whether ξ is $\mathcal{U}$ or $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$.

Very roughly speaking, this is done as follows: $\mathcal{C}$ partitions its inputs into t blocks of size m. For each block $B_i = (b_{i1}, \ldots b_{im})$, $\mathcal{C}$ "acts as if" the inputs in this block form a public key: $\mathcal{C}$ generates a block of random encryptions of zero and a block of random encryptions of one under this hypothetical public key and feeds B_i followed by these two blocks of encryptions to $\mathcal{A}$ (the blocks are ordered at random). $\mathcal{A}$ responds with a guess of which block is first. If $\mathcal{A}$ is correct sufficiently frequently, then $\mathcal{C}$ concludes that $\xi = \mathcal{H}_{u,n^{-\mathcal{D}_1},n}$; otherwise $\mathcal{C}$ concludes that $\xi = \mathcal{U}$. The intuition is that if $\xi = \mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ then each block B_i is a valid public key, so by assumption $\mathcal{A}$ has a non-negligible probability of distinguishing encryptions of zero from encryptions of one. On the other hand, if $\xi = \mathcal{U}$ then all the "encryptions" of zero that $\mathcal{C}$ generates are just sums of uniformly distributed random vectors; hence, $\mathcal{A}$ would have to distinguish between two almost identical distributions, which is impossible.

For the rest of the proof of Theorem 7.1, suppose we are given a basis of a lattice L whose shortest vector is unique up to a factor of $n^{\mathcal{D}_2}$. Let v be a shortest non-zero vector in L. Let X be the set of all $u \in \mathbb{R}^n$ so that $\frac{1}{2} \leq \|u\| \leq 1$ and $\mathcal{C}$ distinguishes the random variables $\mathcal{U}$ and $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ with probability at least $\frac{1}{2} + n^{-c_2}$. We describe a probabilistic polynomial time machine $\mathcal{B}$ that finds v.

$\mathcal{B}$ generates a number t of linear transformations $U_1, \ldots, U_t$, where each U_i can be written $U_i = \theta\nu$ where $\theta \in \mathbb{R}$ and ν is an orthogonal linear transformation. Intuitively, ν rotates the lattice L leaving the lengths of the basis vectors unchanged, while θ scales the rotated basis. We argue (in the full paper) that with probability at least $1-2^{-2n}$, at least one of the vectors $U_i v$ is in X. Fix such an i; we will find $U_i v$, the shortest vector of $U_i L$. Since $U_i = \theta\nu$, we have that w is an $n^{\mathcal{D}_2}$-unique shortest vector of L iff $U_i w$ is an $n^{\mathcal{D}_2}$-unique shortest vector of $U_i L$. Moreover, since $U_i v \in X$ we have $\frac{1}{2} \leq \|U_i v\| \leq 1$ and any vector in $U_i L$ not parallel to $U_i v$ has length at least $\frac{1}{2} n^{\mathcal{D}_2} > 1$. It follows that J, the dual lattice of $U_i L$, is a *random* $(n^{\mathcal{D}_2}, 1)$ lattice. It is random because U_i is random.

$\mathcal{B}$ chooses a new system of coordinates so that $U_i e_j$, $j = 1, ..., n$ is the new basis. Let $K = \mathcal{K}$, that is, $KU^{(n)} = \mathcal{Q}$. We prove that the distance of the distributions of $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ and $\xi_{J,K,R}$ is exponentially small; moreover, clearly $\eta_K = \mathcal{U}$. Therefore the distinguishability of $\mathcal{U}$, and $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ would imply the distinguishability of $\xi_{J,K,R}$ and η_K; hence, as argued in the case of the second cryptosystem, there is a probabilistic polynomial time algorithm to find $J^{(d,M)}$, and so the shortest vector of $U_i L$, using $\mathcal{A}$ as an oracle, with a probability exponentially close to one.

References

[1] L. Adleman, On Breaking Generalized Knapsack Public Key Cryptosystems, Proceedings 15th Annual ACM Symposium on Theory of Computing, 1983, pp. 402–412

[2] M. Ajtai, Generating Hard Instances of Lattice Problems, Proceedings 28th Annual ACM Symposium on Theory of Computing, 1996

[3] M. Ajtai, C. Dwork, Lattice Based Cryptography, submitted to STOC97 (Appendix 1)

[4] M. Ajtai and R. Fagin, Reachability is Harder for Directed than for Unidirected Graphs, *J. Symbolic Logic 55*(1), pp. 113 – 150, 1990

[5] D. Boneh and R. Venkatesan, Hardness of Computing the Most Significant Bits of Secret Keys in Diffie-Hellman and Related Schemes, CRYPTO'96, 1996

[6] J.W.S. Cassels, *An Introduction to the Geometry of Numbers*, Springer, 1959

[7] D. Coppersmith, Finding a Small Root of a Univariate Modular Equation, *Eurocrypt'96*, 1996

[8] W. Diffie, The First Ten Years of Public Key Cryptography, . *Proc. IEEE 76*(5), pp. 560-577, 1988

[9] W. Diffie and M.E. Hellman, New Directions in Cryptography, *IEEE Transactions on Information Theory*, v.IT-22, n.6, pp. 644–654, 1976

[10] D. Gordon, Discrete Logarithms in $GF(P)$ Using the Number Field Sieve, SIAM J. Discrete Mathematics, pp. 124–138, 1993

[11] O. Goldreich, *Lecture Notes on Foundations of Cryptography*, http://www.wisdom.weizmann.ac.il/people/homepates/oded/ln89.html, 1989 (see also, *Foundations of Cryptography (Fragments of a Book)*, http://www.wisdom.weizmann.ac.il/people/homepates/oded/frag.html)

[12] P.M. Gruber, C.G.Lekkerkerker, *Geometry of Numbers*, North-Holland, 1987

[13] M. Grötschel, L. Lovász, A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer, Algorithms and Combinatorics 2, 1988

[14] J. Hastad, Solving Simultaneous Modular Equations of Low Degree, *SIAM J. Computing 17*(2), pp.336–341, 1988

[15] J. Hastad, J. Lagarias, A. Frieze, *SIAM J. Computing 17*(3), 1988

[16] J.C. Lagarias, A.M. Odlyzko, Solving low-density subset sum problems, *Journal of the Association for Computing Machinery 32* pp. 229-246, 1985. An earlier version appeared in *Proc. 24th Annual Symposium on Foundations of Computer Science*, 1983.

[17] R. Rivest, A. Shamir, L. Adelman, A Method for Obtaining Digital Signatures and Public-Key Cryptosystems, *CACM 21*(2), pp. 120–126, 1978

[18] A. Shamir, A Polynomial-Time Algorithm for Breaking the Basic Merkle-Hellman Cryptosystem, *Proc. 23rd Annual Symposium on Foundations of Computer Science*, pp. 145–152, 1982

Appendix 1
Lattice-Based Cryptography

Miklos Ajtai and Cynthia Dwork

We present a probabilistic public key cryptosystem based on a problem involving lattices.

1 Introduction

Historically, public-key cryptosystems have almost without exception been built on the assumed hardness of solving random instances of trapdoor knapsack systems, computing discrete logarithms modulo a prime, and factoring [8]. Among knapsack systems, all but the Chor-Rivest schemes have been broken (*e.g.*, [1, 16, 18]). The narrowness of the remaining options has been cited as a potential fragility of public-key cryptography [8].

We construct a probabilistic public key cryptosystem whose security is based on the assumption that it is computationally infeasible to find the shortest vector in a random instance of a certain class of lattices in which the shortest vector is unique in a sense described below. In building the cryptosystem we actually work with the duals of lattices in this class. Finding the shortest vector in the lattice is equivalent to finding a certain hyperplane in the dual. For this reason we say the cryptosystem is based on the *hidden hyperplane* assumption.

Our cryptosystem differs from previous constructions in that it does not depend on a trapdoor function, but rather on some probabilistic properties. Our approach also yields a conceptually simple and extremely natural pseudo-random generator.

1.1 Related Work

Lattices, and in particular lattice basis reduction techniques, have most often been used in cryptography to obtain negative results: breaking knapsack cryptosystems [1, 16, 18], breaking linear congruantial pseudo-random generators [15], breaking broadcast schemes relying on low-exponent RSA [14], breaking random padding of low-exponent RSA messages [7].

In a recent paper, a positive result on the hardness of computing the most significant bits of a Diffie-Hellman secret key was obtained by rounding in lattices using basis reduction [5]. Also recently, Ajtai [2] obtained a technique for generating solved random instances of a class of problems involving lattices, with the property that solving a random instance in the class is as hard as solving the hardest instance. In particular, he showed that if there is a probabilistic polynomial time algorithm which finds a short vector in a random lattice with probability at least $\frac{1}{2}$, then there is an algorithm which solves the following problem (among other famous problems):

(P2) Find the shortest nonzero vector in an n-dimensional lattice L where the shortest vector v is unique, in the sense that any other vector whose length is at most $n^c\|v\|$ is parallel to v, where c is a sufficiently large absolute constant.

Lattices in which the shortest vector is unique in a sense similar to that of (P2) play an important role in attacking knapsack cryptosystems (see [16]). (The polynomial factor of (P2) is substituted by an exponential one.) The difficulty of breaking our cryptosystem is related to the difficulty of solving random instances of (P2). We describe this in more detail below.

1.2 Hidden Hyperplanes

If $a_1, \ldots, a_n$ are linearly independent vectors in $\mathbb{R}^n$, then we say that the set $L(a_1, \ldots, a_n) = \{\sum_{i=1}^n \lambda_i a_i \mid a_1, \ldots a_n \in \mathbb{Z}\}$ is a lattice in $\mathbb{R}^n$. The set $a_1, \ldots, a_n$ is called a basis of the lattice. The length of the basis is the length of the longest basis vector.

Assume n is a positive integer, $M > 0$, $d > 0$ are real numbers, and $L \subseteq \mathbb{Z}^n$ is a lattice which has an $n-1$ dimensional sublattice L' with the following properties:

1. L' has a basis of length at most M;

2. if H is the $n-1$ dimensional subspace of $\mathbb{R}^n$ containing L' and $H' \neq H$ is a coset of H intersecting L, then the distance of H and H' is at least d.

Then we say that L is a (d, M)-lattice. If $d > M$, then L' is unique. In this case L' will be denoted by $L^{(d,M)}$. The minimum distance between H and a coset of H intersecting L will be denoted d_L.

Let $c > 5$ be a real number, and let $\mathcal{L}$ be a distribution on the set of (d, M) lattices for which $d > n^c M$ and $d \leq d_L \leq 2d$. The *hidden hyperplane* assumption for $\mathcal{L}$ says that, given a random (d, M) lattice $L \in_R \mathcal{L}$, it is computationally infeasible to compute $L^{(d,M)}$.

The hidden hyperplane assumption is related to Problem (P2) as follows: if L is a (d, M) lattice, then L^*, the dual of L, has a shortest vector v of length $1/d_L \leq 1/d$, and this vector is unique up to a factor of n^c. The vector v is orthogonal to H, the $n-1$ dimensional subspace of $\mathbb{R}^n$ containing $L^{(d,M)}$, and given $L^{(d,M)}$ it is possible to find v. Thus, given a (d, M) lattice Λ, a solution to the hidden hyperplane problem for Λ^* yields the unique shortest vector in Λ.

We first consider only the case in which the distribution $\mathcal{L}$ on (d, M) lattices is constrained so that the lattices are presented by a random basis whose length is greater than d_L by only a polynomial (in n) factor. In Section 5 we remove this constraint.

1.3 Design of the Cryptosystem

We base the cryptosystem on the hidden hyperplane assumption, so that the ability, given a random basis B for $L \in_R \mathcal{L}$, to distinguish encryptions of zeros from encryptions of ones with polynomial probability implies the ability to find $L^{(d,M)}$. Roughly speaking, the public key will be a random basis of a (d, M) lattice $L \in_R \mathcal{L}$ for a suitably chosen distribution $\mathcal{L}$. The corresponding private key will be a basis for $L^{(d,M)}$. An encryption of 0 will be a slightly perturbed lattice point in $\mathbb{R}^n$, and an encryption of 1 will be a random point in $\mathbb{R}^n$. We prove (Theorem 3.1) that the ability to distinguish these two distributions with polynomial advantage yields a probabilistic algorithm that with all but exponentially small probability finds $L^{(d,M)}$.

The Key Pair Generation Procedure

1. Generate a random $n-1$ dimensional lattice L' which has a basis $b_1, \ldots, b_{n-1}$ such that $\|b_i\| \leq M$; for example, we can use the random class given in [2]. Let H be the $n-1$ dimensional subspace containing L'.
2. Choose $d \geq n^c M$.
3. Choose a random vector b_n of distance $d \leq d_L \leq 2d$ from H.
4. The private key is any basis of H.
5. Construct a random basis B' for $L = L(B)$. The public key is (B', M).

The Encryption and Decryption Procedures

Roughly speaking, the encryption protocol is as follows. To encrypt 0, choose a random lattice point in the cube KU^n, where U^n is the n dimensional unit cube and $K \geq 2^n d$. In addition, choose $m \geq c_0 n$ random vectors in the n dimensional ball of radius $n^3 M$ centered at the origin. We refer to the sum of the m vectors as the *perturbation*. The ciphertext is the sum of the lattice point and the perturbation. To encrypt 1, choose a random (probably non-lattice) point in KU^n; this point is the ciphertext.

Given a ciphertext z, the receiver first computes the distance of the ciphertext from the nearest coset of H intersecting L. If the distance is sufficiently small, then z is interpreted as an encryption of 0; otherwise, z is interpreted as an encryption of 1.

We actually describe two versions of the cryptosystem. The scheme described in Section 5 requires that the lattice L be presented by a basis that is of length at most $n^c d_L$ for some constant c. We call this the *constrained* version. This restriction is relaxed in Section 5 (the *unconstrained* version).

The remainder of the paper is organized as follows. Section 2 contains definitions and a small amount of background material. Section 5 proves the indistinguishability reduction for the constrained version. Section 4 gives precise descriptions of the key generation, encryption, and decryption procedures to fit into the framework of Section 5. Section 5 extends the results of Sections 5 and 4 to the unconstrained version of the scheme.

2 Definitions

The fundamental concepts concerning lattices and public-key crytposystems can be found in [6, 11, 12, 13, 9, 17].

A *lattice* in $\mathbb{R}^n$ is a set of the form

$$L = L(b_1, \ldots, b_n) = \{\sum_{i=1}^{n} \lambda_i b_i \mid \lambda_i \in \mathbb{Z}, i = 1, \ldots, n\},$$

where $b_1, \ldots, b_n$ is a basis of $\mathbb{R}^n$. We say that $(b_1, \ldots, b_n)$ is a *basis* of L. The *length* of a vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, denoted $\|x\|$ is $(x_1^2 + \ldots + x_n^2)$. The length of the basis $(b_1, \ldots, b_n)$ is the length of the longest basis vector, $\max_{i=1}^{n} \|b_i\|$. The *determinant* of L, denoted $\det(L)$, is the absolute value of the determinant of the parallelepiped with sides $b_1, \ldots, b_n$, where $b_1, \ldots, b_n$ is *any* basis for the lattice: $\det(L) = |\det(b_1, \ldots, b_n)|$.

The *dual* lattice of L, denoted L^*, is defined as

$$L^* = \{x \in \mathbb{R}^n \mid x^T y \in \mathbb{Z} \text{ for all } y \in L\}.$$

If $(b_1, \ldots, b_n)$ is a basis of L then $(c_1, \ldots, c_n)$ is a basis for L^*, where

$$c_i^T b_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

If $a_1, \ldots, a_n \in \mathbb{R}^n$ are linearly independent vectors, then $\mathcal{P}^-(a_1, \ldots, a_n)$ denotes the half-closed parallelepiped

$$\{\sum_{i=1}^{n} \gamma_i a_i \mid 0 \leq \gamma_i < 1, i = 1, \ldots, n\}.$$

Let $\mathcal{P}' = \mathcal{P}^-(a_1, \ldots, a_n)$. By "reduced modulo $\mathcal{P}'$" we mean reduced modulo $(a_1, \ldots, a_n)$ into the lattice parallelepiped $\mathcal{P}'$.

Let P_1 and P_2 denote two probability distributions and let Ω be a σ-field. The *distance* between P_1 and P_2 is

$$\sup\{|P_1(A) - P_2(A)| + |P_1(B) - P_2(B)| \text{ s.t. } A, B \in \Omega,\ A \cap B = \emptyset\}.$$

The n *dimensional ball of radius* R is the set of vectors $x \in \mathbb{R}^n$ such that $\|x\| \leq R$.

3 Reduction for the Constrained Version

In this section we prove that if we constrain the distribution $\mathcal{L}$ so that each (d, M) lattice $L \in \mathcal{L}$ can be presented by a basis whose length exceeds d_L by at most a polynomial (in n) factor, then the ability to distinguish encryptions of 0's from encryptions of 1's yields the ability to solve the hidden hyperplane problem. This implies that the only way to break the crytosystem is to find the private key.

We first describe a procedure from [2] for sampling lattice points within a small cube. This procedure is used in the proof of the reduction for the constrained version (Section 3.2), and in generating the public key (Section 4.1). A variation, used for sampling lattice points within an exponential (in n) sized cube, is used in the reduction and in generating ciphertexts for both schemes.

3.1 Sampling Lattice Points Within a Cube

Lemma 3.1 *There exists a real $c_1 > 0$ such that for all real $c \geq c_1$ there exists a real $c_2 > 0$ such that there is a polynomial time procedure that, given a lattice $L \subseteq \mathbb{Z}^n$ represented by a basis $Y = (Y_1, \ldots, Y_n)$, chooses random lattice points inside a cube in $\mathbb{R}^n$ whose sides have length $n^c\|Y\|$, with a distribution whose distance from the uniform distribution is at most $2^{-n^{c_2}}$.*

Remark 3.1 *It is also possible to give such a procedure which defines a uniform distribution on n^cY, using the fact that the points of a lattice parallelepiped form a finite Abelian group, whose order can be computed. Thus a uniform distribution can be defined on them by using a system of generators.*

Proof: At a high level, the lattice points are obtained as follows. We find a parallelepiped $\mathcal{P}$ whose vertices are lattice points, such that the cube is contained within $\mathcal{P}$ and the number of lattice points contained in $\mathcal{P}$ is at most a polynomial times the number of lattice points contained in the cube. We then choose lattice points randomly and (almost) uniformly inside of $\mathcal{P}$, discarding those outside of the cube.

Finding $\mathcal{P}$

We find the lattice parallelepiped $\mathcal{P}$ as done in Lemma 3 of [2]. Let $y = \max_{i=1}^n \|Y_i\|$. Let $M' = n^c y$, for a fixed c to be chosen later. Let $f_1, \ldots, f_n$ be pairwise orthogonal vectors of length n^3M'. For each i, let b_i be a lattice vector so that $\|f_i - b_i\| \leq \frac{1}{2}nM'$. Let $Q = \mathcal{P}(f_1, \ldots, f_n)$, $\mathcal{P} = \mathcal{P}(b_1, \ldots, b_n)$, and let Q' be the cube obtained by shrinking Q from its center by a factor of $1 + \frac{1}{2n}$.

$Q' \subseteq \mathcal{P}$ will be the cube in the statement of the lemma. To argue that the number of lattice points in Q' is polynomially related to the number of lattice points in $\mathcal{P}$, we first obtain an upper bound on the volume of $\mathcal{P}$ and then apply a lemma of [2] relating the number of lattice points in a parallelepiped $\mathcal{P}$ to the volume of $\mathcal{P}$ and the determinant of the lattice.

Let Q'' be obtained by enlarging Q from its center by a factor of $1 + \frac{1}{2n}$. Then $Q' \subseteq \mathcal{P} \subseteq Q''$, so volume($Q'$) $\leq$ volume($\mathcal{P}$) $\leq$ volume(Q''), and, since $\frac{1}{2} \leq (1 + \frac{1}{2n})^{-n}$ and $(1 + \frac{1}{2n})^n \leq 2$ we have $\frac{1}{2}(n^3M')^n \leq$ volume($\mathcal{P}$) $\leq 2(n^3M')^n$.

Ratio of Numbers of Lattice Points in Q' and $\mathcal{P}$

The next lemma says that if a parallelepiped is not too skewed, then the number of lattice points contained in it is proportional to its volume. Applying the lemma to the cubes Q' and Q'' yields a lower bound on the ratio $|L \cap Q'|/|L \cap Q''| \leq |L \cap Q'|/|L \cap \mathcal{P}|$.

Lemma 3.2 (Ajtai [2]]: *Assume that $L = L(a_1, \ldots, a_n)$ is a lattice in $\mathbb{R}^n$, where $\|a_i\| \le y$, $i = 1, \ldots, n$ and $g_1, \ldots g_n$ are linearly independent vectors in $\mathbb{R}^n$ (not necessarily in L) and $b \in \mathbb{R}^n$. Let k_0, respectively k_1, be the number of lattice points in the closed set $b + \mathcal{P}(g_1, \ldots, g_n)$, respectively in its interior. Let H be the minimal height, let V be the volume, and let S be the surface area of $\mathcal{P}(g_1, \ldots, g_n)$. Then*

(a) $(detL)^{-1}(1 - \frac{2yn}{H})^n V \le k_j \le (detL)^{-1}(1 + \frac{2yn}{H})^n V, \quad j = 0, 1$

(b) If F is a hyperplane then the number of lattice points in $F \cap (b + \mathcal{P}(g_1, \ldots, g_n))$ is at most $2Syn(1 + 2\frac{2yn}{H})^{n-1}(detL)^{-1}$.

■

Let $H_{Q'}$, repectively, $H_{Q''}$, denote the height of Q', respectively, Q''. Then $H_{Q'} = n^3 M'(1+\frac{1}{2n})^{-1}$ and $H_{Q''} = n^3 M'(1+\frac{1}{2n})$, volume$(Q') = (n^3 M'(1+\frac{1}{2n})^{-1})^n$, and volume$(Q'') = (n^3 M'(1 + \frac{1}{2n}))^n$. Applying part (a) twice we obtain

$$1 \ge \frac{|L \cap Q'|}{|L \cap Q''|} \ge \frac{(detL)^{-1}(1 - \frac{2yn}{H_{Q'}})^n \text{volume}(Q')}{(detL)^{-1}(1 + \frac{2yn}{H_{Q''}})^n \text{volume}(Q'')}.$$

It has been argued above that the ratios of the volumes is at least $1/4$. Assuming $n \ge 2$ and $n^c y = M'$ in the expressions for $H_{Q'}$ and $H_{Q''}$, we get $(1 - \frac{2yn(1+\frac{1}{2n})}{n^{3+c}y})^n = (1 - \frac{2(1+\frac{1}{2n})}{n^{c+2}})^n \ge e^{-1}$ in the numerator and $(1 + \frac{2yn}{n^{c+3}y(1+\frac{1}{2n})})^n = (1 + \frac{2}{n^{c+2}(1+\frac{1}{2n})})^n \le e^2$ in the denominator. This bounds the ratio from below by $1/4e^3$.

Sampling Almost Uniformly in $\mathcal{P}$

To choose lattice points almost uniformly from $\mathcal{P}$, we choose lattice points x from a much larger lattice cube and reduce x modulo $\mathcal{P}$, using two additional lemmas from [2], described next.

Lemma 3.3 ([2]): *Assume that $a_1, \ldots, a_n \in \mathbb{R}^n$ are linearly independent. Then for each $b \in \mathbb{R}^n$, there is a unique $b' \in \mathcal{P}^-(a_1, \ldots, a_n)$ so that $b - b' \in L(a_1, \ldots, a_n)$. Moreover, if $b, a_1, \ldots a_n \in \mathbb{Z}^n$, then b' can be computed in polynomial time in* size$(b) + \sum_{i=1}^n$ size(a_i).

Proof: Express b as a linear combination of the vectors a_i and take the integral part of the coefficients. Assume that we get the vector $v = \sum_{i=1}^n r_i a_i$. Then $b' = b - v$ satisfies the requirement. The uniqueness of b' is trivial. ■

We denote the b' obtained in Lemma 3.3 by $b_{(\bmod\ a_1, \ldots, a_n)}$.

Lemma 3.4 ([2]): *For all $d_1 > 0$ there is a $d_2 > 0$ so that the following holds. Assume that $y_1, \ldots, y_n$ are linearly independent vectors in $\mathbb{Z}^n$, $\sigma \ge n$, and $b_1, \ldots, b_n \in L = L(y_1, \ldots, y_n)$ is a set of linearly independent vectors as well, with $\max_{i=1}^n \|b_i\| \le 2^{\sigma^{d_1}}$ and $\max_{i=1}^n \|y_i\| \le 2^{\sigma^{d_1}}$. Suppose further that $\mu_1, \ldots, \mu_n$ are independent random variables which take their values with uniform distribution on the integers in the interval $[0, 2^{\sigma^{d_2}}]$. Let $\chi = (\sum_{i=1}^n \mu_i y_i)_{(\bmod\ b_1, \ldots b_n)}$. Then the distribution of χ on the points of $L \cap \mathcal{P}^-(b_1, \ldots b_n)$ is almost uniform in the following sense:*

if for each $v \in \mathcal{P}^-(b_1, \ldots b_n)$, $p_v = \Pr(\chi = v)$ and k is the number of lattice points in $\mathcal{P}^-(b_1, \ldots b_n)$, then $\sum_{v \in \mathcal{P}^-(b_1, \ldots b_n)} |p_v - \frac{1}{k}| \le 2^{-\sigma^{d_1}}$.

Applying this lemma to a large parallelepiped enables us to sample lattice points (almost) uniformly in the interior of $\mathcal{P}(b_1, \ldots, b_n)$. Specifically, choose d_1 such that $\max_{1 \le i \le n} \|b_i\| \le 2^{n^{d_1}}$ and take c_2 in the statement of Lemma 3.1 to be d_1. The large parallelepiped is $\mathcal{P}(2^{\sigma^{d_2}} y_1, 2^{\sigma^{d_2}} y_2, \ldots, 2^{\sigma^{d_2}} y_n)$. ■

Lemma 3.5 *Given a lattice $L \subseteq \mathbb{Z}^n$ represented by a basis $Y = (Y_1, \ldots, Y_n)$, and given an integer $K \ge 2^{cn}\|Y\|$, $c > 0$, there is a polynomial time procedure that chooses random lattice points inside KU^n with a distribution whose distance from the uniform distribution is at most $2^{-n^{c'}}$ for some $c' > 0$.*

Proof: (Sketch) Let Z be the cube obtained by stretching KU^n from its center by a factor of $(1 + \frac{1}{2n})$. Following the construction in Lemma 3.1 we find a lattice parallelepiped $\mathcal{P}$ whose vertices are close to those of Z. Then KU^n is contained in $\mathcal{P}$ and the number of lattice points in $\mathcal{P}$ exceeds the number of lattice points in KU^n by at most a polynomial factor. We then use Lemma 3.4 to sample KU^n with a distribution exponentially close to uniform. ■

3.2 Indistinguishability of Distributions

We will frequently need to choose a vector t uniformly from $S^n(R)$. We do this inductively, one coordinate at a time, beginning with the nth coordinate. The probability density function describing the choice of the kth coordinate in a k dimensional ball of radius R_k is as follows: the probability of choosing the kth coordinate to have a value of at least $\frac{x}{R_k}$ is $\int_x^1 (\sqrt{1-x^2})^{k-1} dx / \int_0^1 (\sqrt{1-x^2})^{k-1} dx$. We take $R_n = R$. Once the kth coordinate is chosen, say it has value x_k, we recursively choose the remaining coordinates uniformly in the $k-1$ dimensional ball of radius $R_{k-1} = R_k \sqrt{1 - x_k^2}$

Let L be a lattice and let $K > 0, R > 0$ be real numbers. The random variable $\xi_{L,K,R}$ is defined in the following way: first we take a point x chosen uniformly at random from $KU^{(n)} \cap L$, where $U^{(n)}$ is the unit cube in $\mathbb{R}^n$. Then we choose m vectors $t_1, \ldots, t_m$ uniformly at random from $S^n(R)$, where $m = c_0 n$ for some $c_0 \ge 4$. The value of $\xi_{L,K,R}$ is $x + \sum_{i=1}^m t_i$.

η_K will be a random variable whose values are taken with uniform distribution on $KU^{(n)}$. δ will be a random variable taking the values 0 and 1 with probabilities $\frac{1}{2}$, $\frac{1}{2}$. $\nu_{L,K,R}$ is defined in the following way. We randomize $\delta, \xi_{L,K,R}$ and η_K independently. If $\delta = 0$, then $\nu_{L,K,R} = \eta_K$, if $\delta = 1$ then $\nu_{L,K,R} = \xi_{L,K,R}$.

Suppose that the real number $c > 5$ and the positive integers n, d, M, K, R, $d > n^c M$ are given, and $\mathcal{L}$ is a distribution on (d, M)-lattices in $\mathbb{Z}^n$. We say that a probabilistic algorithm $\mathcal{A}$ finds $L^{(d,M)}$ on $\mathcal{L}$ with a probability p, if given as input a description of $\mathcal{L}$ (including d and M) and $L \in \mathcal{L}$, $\mathcal{A}$ outputs $L^{(d,M)}$ with probability p, where the probability is taken both for the randomization of L and for the randomization in $\mathcal{A}$. Sometimes we will allow $\mathcal{A}$ to use an oracle. In this case each use of the oracle will be counted as one time unit in the definition of the time used by $\mathcal{A}$.

We assume a model in which for some constants e_0 and e_1, a $2^{-n^{e_0}}$ approximation to a real input can be obtained in time $\theta(n^{e_1})$. Suppose that the real number $c > 5$ and the positive integers $n, d, M, K, R,\ d > n^c M$ are given, and $\mathcal{L}$ is a distribution on (d, M)-lattices in $\mathbb{Z}^n$. We say that the probabilistic algorithm $\mathcal{A}$ *distinguishes* $\xi_{L,K,R}$ *and* η_K *on* X *with a probability* p if given $L \in_R \mathcal{L}$ and a random value of $\nu_{L,K,R}$ as an input (together with n, M, d, K, R), $\mathcal{A}$ outputs a $0,1$ value w so that $P(w = \delta) = p$. Note that in polynomial time $\mathcal{A}$ sees only polynomial (in n) bits of its input.

Theorem 3.1 *There exist $c, c_4, c_5, c_6 > 0$ so that for all $c_1 > 0,\ c_2 > 0$ there exists $c_3 > 0$ and a probabilistic algorithm $\mathcal{B}$ (using an oracle) so that if n, d, M, K, R are positive integers satisfying the inequalities,*
(1) $\log d + \log M + \log K + \log R < n^{c_1}$
(2) $d > n^{c_6} M$,
(3) $R > n^c M$,
(4) $2^{c_5 n} d > K > 2^{c_4 n} d$,
and $\mathcal{L}$ is a distribution on the set of (d, M) lattices in $\mathbb{Z}^n$ presented by vectors of length at most $n^{c_1} d_L$ and for which $d_L > n^5 M$ and $d \le d_L \le 2d$, and $\mathcal{A}$ is a probabilistic algorithm which distinguishes $\xi_{L,K,R}$ and η_K on $\mathcal{L}$ with probability at least $\frac{1}{2} + n^{-c_2}$, then $\mathcal{B}$, using $\mathcal{A}$ as an oracle, finds $L^{(d,M)}$ on $\mathcal{L}$ with a probability at least $1 - 2^{-n}$, in time n^{c_3}.

Let $L \in_R \mathcal{L}$ be presented by $(b_1, \ldots, b_n)$ such that $\max_{1 \le i \le n} \|b_i\| \le n^{c_1} d_L$. Strictly speaking, as described above, we must charge time $\theta(n^{e_1})$ for $\mathcal{B}$ to access a $2^{-n^{e_0}}$ approximation to a real input. For simplicity, we first describe $\mathcal{B}$ as if it could access any real input in a single step, returning to this issue later.

Algorithm $\mathcal{B}$ works as follows. Let $K' = n^c d_L$. Choose a polynomial (in n) random lattice points $p_1, \ldots, p_{m'} \in K'U^n$ using Lemma 3.1. (This is where we use the assumption that L is presented by a basis of length at most polynomial in n larger than d.) For $1 \le i < j \le m'$, let $a_{ij} = p_i - p_j$.

Note that $K'U^n \cap L$ is intersected by at most n^c cosets H' of H intersecting L, where H is the $n-1$ dimensional subspace containing $L^{(d,M)}$. Let H' be a coset of H whose intersection with $K'U^n \cap L$ is maximal. The number of differences a_{ij} such that p_i and p_j are both in H' is at least $(\frac{1}{n^c})^2(\frac{m'(m'-1)}{2})$, so a polynomial fraction of the a_{ij} are in H. The key idea is to use $\mathcal{A}$ to determine which of the differences a_{ij} are in H. By doing so, if m' is sufficiently large, then, by arguments appearing in [2], $\mathcal{B}$ will find a basis for H among the a_{ij}.

Testing for Containment in H

Let $L^{d,M} = L(b_1, \ldots, b_{n-1})$, where $\max_{1 \le i \le n-1} \|b_i\|$, and let $\mathcal{P}' = \mathcal{P}(b_1, \ldots, b_{n-1})$. Let w be a random variable whose value is $w = \sum_{i=1}^{m} t_i$, where for $1 \le i \le m$, $t_i \in_R S^n(n^3 M)$. We will prove that if $\epsilon > 0$ is arbitrary (it may depend on n) and σ is a strip of width ϵ, not too far from the hyperplane H, then the distribution of the projection of w is almost uniform on the hyperlane, modulo the lattice parallelepiped $\mathcal{P}'$, even with the condition that w is in σ.

Each $v \in \{a_{ij} \mid 1 \leq i < j \leq m'\}$ induces a probability distribution as follows. Let u be a random variable with uniform distribution on $KU^n \cap L$, and let α be a random variable distributed uniformly in $[0,1]$. Define the random variable $\delta_v = u + \alpha v + w$. It follows from the uniformity of the projection of w onto H (modulo $\mathcal{P}'$) that distributions obtained by projecting δ_v and $\xi_{L,K,R}$ onto H are almost uniformly distributed on the projection of KU^n on H, independent of whether or not $v \in H$. Moreover, this is true even if we restrict the distributions to the case in which w lies in a strip σ not too far from H.

If $v \in H$, then $u + \alpha v \in H'$, where H' is the coset of H containing $u \in L$. In this case, if v is not too long, then $u + \alpha v + w$ has essentially the same distribution as $\xi_{L,K,R}$: each depends only on the distance from H of its respective copy of w. If $v \notin H$, then since with all but exponentially small probability u and v are not in the same coset of H, the signed distance of $u + \alpha v$ to the nearest coset of H is uniformly distributed in $(-\frac{d_L}{2}, \frac{d_L}{2}]$; so if v is not too long then δ_v has essentially the same distribution as η_K. Thus, the assumed ability of $\mathcal{A}$ to distinguish $\xi_{L,K,R}$ from η_K reveals whether or not $v \in H$.

Let u_H denote a fixed unit vector orthogonal to H. For $w \in \mathbb{R}$, a *width w strip of $\mathbb{R}^n$ parallel to H* is a set σ of points $p \in \mathbb{R}^n$ such that for some interval I of length w $p \cdot u_H$, the signed distance of p to H, is in I.

Suppose that $R > 0, n > 0$ are fixed. We define a random variable $\xi = \xi^{(n,R)}$, in the following way. We take a random point t of the n-dimensional ball of radius R with uniform distribution. ξ is the last component of t.

Lemma 3.6 *There is a $c > 0$ and an integer n_0 so that if $n \geq n_0$ and R are fixed and $\xi_1, ..., \xi_m$ are independent copies of $\xi^{(n,R)}$ and $\epsilon > 0$, $m \geq n$, then the following holds. Let I be an interval of length ϵ so that $I \subseteq [-\frac{m}{40}R, +\frac{m}{40}R]$. Let $\eta = \sum_{i=1}^{m} \xi_i$ and let G be the event: "there are at least $m/4$ integers i in the interval $[1,m]$ so that $|\xi_i| < (1 - \frac{1}{n})R$". Then the conditional probability of G with the condition $\eta \in I$ is at least $1 - 2^{-cm}$.*

Proof: For each fixed positive integer k and for each real number x, let $\rho_k(x)$ be the probability of $x + \sum_{i=1}^{k} \xi_i \in I$. We prove by induction on k that
(1) the function $\rho_k(x)$ is symmetric to the midpoint of I and monotone decreasing in both directions as we get farther from this midpoint.

Since the distribution of ξ is symmetric to 0, the symmetricity is trivial. Let χ be the density function of ξ. Then for all x we have $\rho_k(x) = \int_{x-R}^{x+R} \chi(y-x)\rho_{k-1}(y)dy$. $\chi(y-x)$ is symmetric to x (as a function of y) and $\rho_{k-1}(y)$ is symmetric to the midpoint of I. Both functions are monotone decreasing as we go away from their point of symmetry.

We use the following general statement about symmetric functions. Suppose that f_0, f_1 are symmetric to 0 and for all x, y, $|x| \leq |y|$ implies $f_i(x) \geq f_i(y)$ for $i = 0, 1$. Then for any $0 < z < w$ we have $\int_{-\infty}^{\infty} f_0(y) f_1(y+w) dy \leq \int_{-\infty}^{\infty} f_0(y) f_1(y+z) dy$, provided that both integrals are finite.

The statement is trivial if both f_0 and f_1 are the characteristic functions of finite intervals. Any other functions with the given properties can be approximated with the sum of such functions, so using the distributivity we get the inequality.

Using (1) we may conclude the proof in the following way. Let $\eta_i = \sum_{j<i} \xi_j$. In the following probabilities always mean probabilities with the condition $\eta \in I$. For each $i = 1, ..., m$ we define a define an event A_i depending only on the values of η_i and ξ_i so that
(2) the conditional probability of A_i with any condition on η_i is at most $\frac{1}{20}$.
Let Y be the event that A_i holds for more than $m/10$ values of i. (2) implies:
(3) the probability of Y is smaller than $2^{-c'm}$ for some absolute constant $c' > 0$ (see, *e.g.*, Corollary 7.1 of [4]). Finally we will show that for any values of $\xi_1, ..., \xi_m$ with $|\xi_i| \le R$, $i = 1, ..., m$, we have that $\neg Y$ implies G. This together with (3) implies the assertion of the lemma.

Definition of A_i. Let α be the midpoint of I. A_i holds if $|\xi_i| > (1 - \frac{1}{n})R$ and at least one of the following conditions are satisfied:

(4) $|\alpha - \eta_{i+1}| > |\alpha - \eta_i|$
(5) $|\alpha - \eta_i| < \frac{2}{3}R$.

We will estimate $P(A_i)$ in cases (4) and (5) separately, using the monotonicity of ρ_{m-i} and the explicit formula for χ. To compute the probability that A_i holds with $|\alpha - \eta_{i+1}| > |\alpha - \eta_i|$, conditioned on $\eta \in I$, we randomize ξ_i, compute the probability that in the remaining $m - i$ steps we get to I, and multiply by p_η, the probability that $\eta \in I$. Since we don't know p_η we instead bound the ratio of two conditional probabilities, both conditioned on $\eta \in I$ (so that the p_η terms cancel).

Suppose $\eta_i < \alpha$ (the case $\eta_i > \alpha$ is analogous). Let J be the interval $[\eta_i - R, \eta_i - R(1 - \frac{1}{n})]$. Then the (conditional) probability of reaching I through J (that is, the first step is in J) is $p_\eta \int_J \chi(\eta_i - y)\rho_{m-i}(y)dy$. Similarly, letting J' be the interval $[\eta_i - R(\frac{1}{n}), \eta_i]$, the conditional probability of reaching I through J' is $p_\eta \int_{J'} \chi(\eta_i - y)\rho_{m-i}(y)dy$.

$$
\begin{aligned}
\frac{p_\eta \int_J \chi\eta_i - y)\rho_{m-i}(y)dy}{p_\eta \int_{J'} \chi(\eta_i - y)\rho_{m-i}(y)dy} &= \frac{\int_J \chi(\eta_i - y)\rho_{m-i}(y)dy}{\int_{J'} \chi(\eta_i - y)\rho_{m-i}(y)dy} \\
&\le \frac{\int_J \chi(\eta_i - y)[\max_{x\in J} \rho_{m-i}(x)]dy}{\int_{J'} \chi(\eta_i - y)[\min_{x\in J'} \rho_{m-i}(x)]dy} \\
&= \frac{\max_{x\in J} \rho_{m-i}(x)}{\min_{x\in J'} \rho_{m-i}(x)} \frac{\int_J \chi(\eta_i - y)dy}{\int_{J'} \chi(\eta_i - y)dy} \\
&\le \frac{\int_J \chi(\eta_i - y)dy}{\int_{J'} \chi(\eta_i - y)dy} \\
&= \frac{\int_{1-\frac{1}{n}}^{1} (\sqrt{1 - x^2})^{n-1}dx}{\int_0^{\frac{1}{n}} (\sqrt{1 - x^2})^{n-1}dx} \\
&< \frac{\frac{1}{n}(\sqrt{1 - (1 - \frac{1}{n})^2})^{n-1}}{\frac{1}{n}(\sqrt{1 - (\frac{1}{n})^2})^{n-1}} \\
&= (\sqrt{\frac{2n - 1}{n^2 - 1}})^{n-1}
\end{aligned}
$$

$$\sim \left(\sqrt{\frac{1}{n}}\right)^{n-1}$$

Thus,

$$p_\eta \int_J \chi(\eta_i - y)\rho_{m-i}(y)dy \quad \leq \quad p_\eta \int_{J'} \chi(\eta_i - y)\rho_{m-i}(y)dy(\sqrt{\frac{1}{n}})^{n-1}.$$

Since $p_\eta \int_{J'} \chi(\eta_i - y)\rho_{m-i}(y)dy$ is a conditional probability it is bounded above by 1, so in this case A_i is clearly bounded by 1/40 (for n sufficiently large), and therefore, allowing also for the case $\eta_i > \alpha$, the conditional probability of A_i in case (4) above is at most 1/20.

The conditional probability of A_i in case (5) above is measured similarly. Let us assume that $\eta_i \in [\alpha - \frac{2}{3}R, \alpha]$ (the case $\eta_i \in [\alpha, \alpha + \frac{2}{3}R]$ is handled analogously). If $\xi_i < 0$ then we are again in case (4), so we need only consider the case in which $\xi \geq 0$. Let J be the interval $[\eta_i + R(1 - \frac{1}{n}), \eta_i + R]$, and let J' be the interval $[\alpha, \alpha + R/n]$. Performing a calculation analogous to that in case (4) we have

$$\begin{aligned}
\frac{p_\eta \int_J \chi(\eta_i - y)\rho_{m-i}(y)dy}{p_\eta \int_{J'} \chi(\eta_i - y)\rho_{m-i}(y)dy} &= \frac{\int_J \chi(\eta_i - y)\rho_{m-i}(y)dy}{\int_{J'} \chi(\eta_i - y)\rho_{m-i}(y)dy} \\
&\leq \frac{\int_J \chi(\eta_i - y)dy}{\int_{J'} \chi(\eta_i - y)dy} \\
&\leq \frac{\int_{1-\frac{1}{n}}^{1} (\sqrt{1-x^2})^{n-1}dx}{\int_{\frac{2}{3}}^{\frac{2}{3}+\frac{1}{n}} (\sqrt{1-x^2})^{n-1}dx} \\
&< \left(\sqrt{\frac{2n-1}{\frac{5}{9}n^2 + \frac{4}{3}n - 1}}\right)^{n-1} \\
&< \left(\sqrt{\frac{4}{n}}\right)^{n-1}
\end{aligned}$$

Thus,

$$p_\eta \int_J \chi(\eta_i - y)\rho_{m-i}(y)dy \quad \leq \quad p_\eta[\int_{J'} \chi(\eta_i - y)\rho_{m-i}(y)dy] \left(\frac{4}{n}\right)^{\frac{n-1}{2}}.$$

Since $p_\eta \int_{J'} \chi(\eta_i - y)\rho_{m-i}(y)dy$ is a conditional probability it is bounded above by 1, so in this case A_i is clearly bounded by 1/40 (for n sufficiently large), and therefore, allowing also for the case $|\alpha - \eta_i| < \frac{2}{3}R$ and $\eta_i > \alpha$, the conditional probability of A_i is at most 1/20.

It remains to show that $\neg Y \wedge \neg G$ is impossible. Intuitively, $\neg Y \wedge \neg G$ would imply that for most of the integers $i \in [1, m]$, η_{i+1} would be closer to α than η_i by at least $(1 - \frac{1}{n})R$, that is, η_i would move toward α with large steps, but according to (5) it would only rarely get close to α.

Let us say ξ_i is *large* if $|\xi_i| \geq (1 - \frac{1}{n})R$; otherwise it is *small.* We say that ξ_i moves *away* from α if $|\eta_i - \alpha| < |\eta_{i+1} - \alpha|$; otherwise it moves *toward* α. Finally, we say η_i is *close* to α if $|\eta_i - \alpha| < \frac{2}{3}R$; otherwise it is *far.* Then we can re-phrase the events as follows:

- G is the event that at least $\frac{m}{4}$ of the ξ_i are small;
- Y is the event that A_i holds for more than $\frac{m}{10}$ values of i.
- A_i is the event that $[(\xi_i$ is large and ξ_i moves away) OR (ξ_i is large and η_i is close)].

Negating, we get:

- $\neg G$ is the event that more than $\frac{3m}{4}$ of the ξ_i are large;
- $\neg Y$ is the event that A_i holds for at most $\frac{m}{10}$ values of i;
- $\neg A_i$ is the event that $[(\xi_i$ is small or ξ_i moves toward α) AND (ξ_i is small or η_i is far from α)]. Equivalently, $\neg A_i$ is the event that $[(\xi_i$ is small) OR (ξ_i moves toward α and η_i is far)].

Let us assume $\neg G \wedge \neg Y$. Focussing first on the more than $\frac{3m}{4}$ values of i for which ξ_i is large (from $\neg G$), we have that for at most $\frac{m}{10}$ of these values for i, either ξ_i moves away from α or η_i is close to α (from $\neg Y$).

We break the sequence $(\eta_1, \xi_1), (\eta_2, \xi_2), \ldots, (\eta_m, \xi_m)$ into intervals so that in the odd intervals η_i is far from α and in the even intervals η_i is close to α. The integers j_x defined below are the endpoints of these intervals.
Definitions:

- $j_0 = 0$;
- for $x \geq 0$, $j_{2x+1} = \min_{i > j_{2x}} \eta_i$ is close to α;
- for $x \geq 0$, $j_{2x+2} = \min_{i > j_{2x+1}} \eta_i$ is far from α.

Let $y = \frac{1}{10}$ and let $z = \frac{3}{4}$. Let $D = |\{i$ s.t. (ξ_i is large and ξ_i moves away from α) or (ξ_i is large and η_i is close to α)$\}|$. By assumption ($\neg Y$), $D \leq my = \frac{m}{10}$. Let $S = |\{i$ s.t. ξ_i is small$\}|$. By assumption ($\neg G$), $S < (1 - z)m = \frac{m}{4}$.

For $0 \leq x$, the *xth far interval* is the interval $[j_{2x}, j_{2x+1} - 1]$ (by definition η_i is far from α for all i in this range). Let k be the number of far intervals. For $0 \leq x \leq k-1$, let β_x denote the fraction $\frac{1}{S}$ times the number of small steps during the xth far interval; more formally,

$$\beta_x = \frac{1}{S} \times |\{i \text{ s.t. } i \in [j_{2x}, j_{2x+1} - 1] \ \wedge \ \xi_i \text{ is small}\}|.$$

For $0 \leq x \leq k - 1$, let let γ_x denote the fraction $\frac{1}{D}$ times the number of large steps away from α during the xth far interval; formally,

$$\gamma_x = \frac{1}{D} \times |\{i \text{ s.t. } i \in [j_{2x}, j_{2x+1} - 1] \ \wedge \ \xi_i \text{ is large}\}|.$$

Finally, let w_0 be a binary variable with value 1 if and only if $0 = \eta_1$ is close α, η_2 is far from α, and ξ_1 is large and moves toward α; in general, for $0 \leq x \leq k-1$, let w_x be the binary variable with value 1 if and only if $\xi_{j_{2x}-1}$ is large and moves toward α (note that by definition $\eta_{j_{2x}-1}$ is close to α, so the w_x are counting large steps toward α that move the walk from close to α to far from α).

Let $q = \frac{\alpha}{(1-\frac{1}{n})R} \leq 2\frac{\alpha}{R}$ (provided $n \geq 2$). Then at most q large steps are needed to walk directly (without interruption) from $0 = \eta_1$ to α. Since $(1 - \frac{1}{n})^{-1} \leq 2$, at most two large steps toward α are required to compensate for a single large step away from α. Moreover, by definition, a large step is at least as large as a small step. Finally, for each large step toward α moving the walk from close to α to far from α, at most one large step back toward α is needed to compensate. Using these facts and the definitions of the β_x, γ_x, and w_x, and letting B denote the number of i such that η_i is far, ξ_i is large, and ξ_i moves toward α, (*i.e.*, the number of large steps taken toward α during the far intervals), we get

$$B \;\leq\; q + \sum_{0\leq x\leq k-1} (2\gamma_x D + \beta_x S) + \sum_{0\leq x\leq k-1} w_x. \tag{1}$$

Note that if $\sum_{0\leq x\leq k-1} w_x > ym$ then for strictly more than $ym = \frac{m}{10}$ values of i we have case (5) of A_i (η_i is near α and ξ_i is large), so it follows from the assumption $(\neg Y)$ that $\sum_{0\leq x\leq k-1} w_x \leq ym$. By definition, $\sum_{0\leq x\leq k-1}(2\gamma_x D+\beta_x S) = 2D+S$, so $B \leq \frac{2\alpha}{R}+2D+S+ym$. However, from $(\neg Y \wedge \neg G)$ we have that $B \geq (z-y)m$. Combining this with the bound on B in (1), we get $\frac{2\alpha}{R} + 4ym + m > 2zm$ which is false for the values we have chosen ($\alpha \leq \frac{Rm}{40}, y = \frac{1}{10}, z = \frac{3}{4}$). ∎

Corrolary 3.1 *Let $w = \sum_{i=1}^{m} t_i$, where each t_i is chosen at random from $S^n(R)$. There is a $c > 0$ and n_0 so that if $n \geq n_0$ and R are fixed, $\epsilon > 0$, and I is an interval of length ϵ contained in $[-\frac{mR}{40}, \frac{mR}{40}]$, then the following holds. For $1 \leq i \leq m$, let ξ_i be the signed distance from H to t_i. Let us assume a coordinate system in which we can write $t_i = r_i + s_i$, where $r_i = (0, 0, \ldots, 0, \xi_i)$. Let G be the event "there are at least $\frac{m}{4}$ integers i in $[1, m]$ such that s_i is chosen from an $n-1$ dimensional ball of radius at least n^2M." Let $\eta = \sum_{i=1}^{m} \xi_i$. Then the conditional probability of G with the condition $\eta \in I$ is at least $1 - 2^{-cm}$.*

Lemma 3.7 *Let $L = L(b_1, \ldots, b_n) \in_R \mathcal{L}$, where $L^{(d,M)} = L(b_1, \ldots, b_{n-1})$. Let $m = c_0 n$, for some $c_0 \geq 4$, let $R = n^3 M$, and let $\mathcal{P}' = \mathcal{P}^-(b_1, \ldots, b_{n-1})$. For $n' \geq n$, assume each of $s_1, \ldots, s_{n'}$ is chosen from an $n-1$ dimensional ball of radius at least n^2M. For each s_i, let G_i be the hyperplane parallel to H containing s_i. If we tile each G_i with copies of $\mathcal{P}'$, then with probability at least $1-2^{-n'\log n}$ at least one of $s_1, \ldots, s_{n'}$ is chosen from a tile completely contained in its respective ball.*

Proof: Recall that $\|b_i\| \leq M$, $1 \leq n-1$. Let S_i be the ball from which s_i is chosen. Consider a particular G_i, tiled with copies of $\mathcal{P}'$, and partition the copies into two sets, those lying entirely within S_i and those lying only partially inside S_i. For any vector s_i, let $\mathrm{ppd}(s_i)$ denote the copy of $\mathcal{P}'$ containing s_i. Let C (respectively, D) be the set of vectors s_i such that

$\mathrm{ppd}(s_i)$ lies entirely (respectively, partially) within S_i. Then $\sum_{i=1}^{n'} s_i = \sum_{s_i \in C} s_i + \sum_{s_j \in D} s_j$. For each i, $\Pr[s_i \in C] \leq \frac{1}{n}$, so with probability at least $1 - 2^{-n' \log n}$ at least one of $s_1, \ldots s_{n'}$ is in C. ■

Using the notation in Corollary 3.1, let us say that $w = \sum_{i=1}^{n}(r_i + s_i)$ is *bad* if either fewer than $\frac{m}{4}$ of the s_i are chosen from balls of radius less than $n^2 M$, or at least $\frac{m}{4}$ of the s_i are chosen from balls of large radius but none of these s_i is chosen from a tile completely contained in its respective ball. If it is not bad, then w is *good.*

Lemma 3.8 *Let $L = L(b_1, \ldots, b_n) \in_R \mathcal{L}$, where $L^{(d,M)} = L(b_1, \ldots, b_{n-1})$. Let $m = c_0 n$, for some $c_0 \geq 4$, and let $R = n^3 M$. Let $\mathcal{P}' = \mathcal{P}^-(b_1, \ldots, b_{n-1})$. For $\epsilon > 0$ and for any interval I of length ϵ contained in $[-\frac{mR}{40}, \frac{mR}{40}]$, let $w = \sum_{i=1}^{m} t_i$ where each $t_i \in_R S^n(R)$. Then the distribution obtained by projecting w onto H modulo $\mathcal{P}'$ differs from the uniform distribution on $\mathcal{P}'$ by at most $2^{-c_1 m}$ for some $c_1 > 0$, even with the condition that the signed distance of w to H is in I.*

Proof: By Corollary 3.1, for some $c > 0$ with probability at least $1 - 2^{-cm}$ for at least $\frac{m}{4}$ of the values of i in $[1, m]$, s_i is chosen from an $n-1$ dimensional ball of radius at least $n^2 M$. By Lemma 3.7, with probability at least $1 - 2^{-\frac{m}{4} \log n}$ at least one of these is chosen from a copy of $\mathcal{P}'$ completely contained in its ball. This s_i is uniformly distributed in $\mathcal{P}'$, and therefore so is the projection of $\sum_{i=1}^{m} t_i$.

For integer $q > 0$, let us tile $\mathcal{P}'$ with small parallelepipeds, each of the form $(\sum_{i=1}^{n-1} \frac{n_i}{q} b_i) + \frac{1}{q}\mathcal{P}'$, where $0 \leq n_i < q$, $i = 1, \ldots, n-1$ is a sequence of integers. Let Ω be the set of sets of tiles. For any given tile a, let $P(a)$ denote the probability that the projection of w is in a and let *unif*(a) denote the probability that a point chosen uniformly from $\mathcal{P}'$ is in a. The distance of the distribution of the projection of w from the uniform distribution on $\mathcal{P}'$ is

$$\sup\{|P(A) - \mathit{unif}(A)| + |P(B) - \mathit{unif}(B)| \text{ s.t. } A, B \in \Omega,\ A \cap B = \emptyset\}.$$

Let A denote the subset of tiles in which P is greater than or equal to *unif* and B denote the subset of tiles in which *unif* exceeds P. Then for some $c_3 > 0$ the first term in the expression yields

$$\begin{aligned}
&\sum_{a \in A} P(a) - \mathit{unif}(a) \\
&= \sum_{a \in A} [P(a|w \text{ is good}) \Pr(w \text{ is good}) + P(a|w \text{ is bad}) \Pr(w \text{ is bad}) - \mathit{unif}(a)] \\
&= \sum_{a \in A} [\mathit{unif}(a)((1 - 2^{-c_3 m}) - 1) + P(a|w \text{ is bad}) \Pr(w \text{ is bad})] \\
&= \sum_{a \in A} [\mathit{unif}(a)(-2^{-c_3 m}) + P(a|w \text{ is bad}) \Pr(w \text{ is bad})] \\
&\leq \sum_{a \in A} [\mathit{unif}(a)(-2^{-c_3 m})] + 2^{-c_3 m} \leq 2^{-c_3 m}.
\end{aligned}$$

The analysis for the second term is identical. The difference between the two distributions is therefore at most $2 \cdot 2^{-c_3 m} \leq 2^{-c_1 m}$ for some $c_1 > 0$. ■

Corrolary 3.2 *Let $L = L(b_1, \ldots, b_n) \in_R \mathcal{L}$, where $L^{(d,M)} = L(b_1, \ldots, b_{n-1})$. Let $m = c_0 n$, for some $c_0 \geq 4$, let $R = n^3 M$, and let $\mathcal{P}' = \mathcal{P}^-(b_1, \ldots, b_{n-1})$. Let $\epsilon > 0$ and let I be an interval of length ϵ contained in $[-\frac{mR}{40}, \frac{mR}{40}]$. Consider the distribution D obtained by choosing $p \in_R H$ and projecting $p + w$, where $w = \sum_{i=1}^m t_i$ and each $t_i \in_R S^n(R)$, onto H. Then the distance of D from the uniform distribution on H is at most $2^{-c_1 m}$, for some $c_1 > 0$, even conditioned on the signed distance of w from H being in I. Moreover, if H' is any coset of H intersecting L, then the lemma holds for any distribution for p in which, if we tile H with copies of $\mathcal{P}'$, each copy of $\mathcal{P}'$ is equally likely to contain p.*

Proof: By Lemma 3.8, the distribution obtained by projecting w onto H modulo $\mathcal{P}'$ is within $2^{-c_1 m}$ of the uniform By randomizing the starting point p we eliminate the need to reduce modulo $\mathcal{P}'$: tile H with copies of $\mathcal{P}'$. The probability that p is in any particular copy of $\mathcal{P}'$ is the same. Note that we get almost uniformity of the projection of $p + w$ modulo $\mathcal{P}'$ because of the properties of w which is independent of whether or not $p \in L$. ∎

In order to apply the lemma to the distribution $\xi_{L,K,R}$ we need to bound the ratio of lattice points in KU^n near the surface of the cube.

Lemma 3.9 *Let $B = 2K^{1/c'}$ for some $c' > 1$. Let $C = (\frac{K}{2}, \frac{K}{2}, \ldots, \frac{K}{2})$. Let Q be the cube centered at C with sides of length $K - 2B$. Then*
(1) the ratio of the volume of Q over the volume of KU^n is at least $1 - 2^{-cn}$ for some $c > 0$ and
(2) the ratio of the number of lattice points in the closed set Q to the number of lattice points in the closed set KU^n is at least 2^{-cn} for some constant $c > 0$.

Proof: By definition the ratio of the volume of Q over the volume of KU^n is at least $(1 - \frac{2B}{K})^n$. From the binomial expansion we have $1 - n\frac{2B}{K} < (1 - \frac{2B}{K})^n$, so (1) follows from the fact that $K \geq 2^{c_4 n} d_L$.

By Lemma 3.2, the ratio of the numbers of lattice points is at least $(1 - \frac{2yn}{K-2B})^n/(1 + \frac{2yn}{K})^n \leq 1$, where $y = \max_{i=1}^n \|b_i\|$. Let $\alpha = \frac{2yn}{K-2B}$ and let $\beta = \frac{2yn}{K}$. Since $\frac{1-\alpha}{1+\beta} > 1 - (\alpha + \beta)$ and $1 - n(\alpha + \beta) \leq (1 - (\alpha + \beta))^n$, the ratio differs from 1 by at most $n(\alpha + \beta) \leq \frac{4yn^2}{K-2B}$. Since $y \leq d_L$ and $K \geq 2^{c_4 n} d_L$ for some $c_4 > 0$, this quantity is exponentially small in n. ∎

Lemma 3.10 *Let $L = L(b_1, \ldots, b_n) \in_R \mathcal{L}$, where $L^{(d,M)} = L(b_1, \ldots, b_{n-1})$. Let $v \in L$ satisfy $|v| \leq K^{1/c}$ for some $c' > 1$. Then:*

1. *if $v \in H$ then the distance between the distributions $\xi_{L,K,R}$ and δ_v is at most 2^{-cn} for some $c > 0$;*

2. *if $v \notin H$ then the distance between the distributions η_K and δ_v is at most 2^{-cn} for some $c > 0$.*

Proof: Let $\mathcal{P}' = \mathcal{P}^-(b_1, \ldots, b_{n-1})$. Let t be the random variable obtained by choosing uniformly a vector in $S^n(n^3M)$. Let w be the random variable defined by $w = \sum_{1\leq i\leq m} t_i$, where each t_i is a copy of t. Let us write $\delta_v = u_0 + \alpha v + w_0$ and $\xi_{L,K,R} = u_1 + w_1$ where $u_0, u_1 \in_R KU^n \cap R$ and w_0, w_1 are copies of w.

Proof of (1): Suppose that, rather than being drawn from $\mathcal{L} \cap KU^n$, we were to choose $u \in_R L$. Fix any coset H' of H intersecting L, and tile H' with copies of $\mathcal{P}'$. Then $u_0 + \alpha v$ is uniformly chosen from among the tiles, as is u_1. For $\epsilon > 0$, let us partition the space $\mathbb{R}^n$ into strips of width ϵ parallel to H. For each strip σ, the difference between the projection of w_0 onto H and the uniform distribution on H is exponentially small in $m \log m$, even conditioned on w_0 being in σ. The same applies to w_1 so the difference between these two distributions is exponentially small in $m \log m$.

Let $\delta'_v = u_0 + \alpha v + w_0$ and $\xi = u_1 + w_1$ where both u_0 and u_1 are chosen unformly from L. Let $\mathcal{T}$ be the parallelepiped $\mathcal{T} = \mathcal{P}^-(\frac{1}{q}b_1, \ldots, \frac{1}{q}b_{n-1}, u_\epsilon)$, where u_ϵ is a vector of length ϵ perpendicular to H. We tile each strip σ with copies of $\mathcal{T}$. Let Ω be the set of sets of tiles. Let A be the set of tiles T for which $\Pr(\xi \in T) > \Pr(\delta_v \in T)$. (The analysis for the set of tiles in which the opposite holds is analogous.) We specify a strip $\sigma = (H_\sigma, I_\sigma)$ by naming H_σ, the coset of H intersecting L nearest to σ, and I_σ, the interval containing the distances of the points in σ from H_σ.

$$\begin{aligned}
& \sum_\sigma \sum_{T\in A\cap\sigma} \left(\Pr[\xi \in T|\xi \in \sigma]\Pr[\xi \in \sigma] - \Pr[\delta'_v \in T|\delta'_v \in \sigma]\Pr[\delta_v \in \sigma]\right) \\
= \;& \sum_\sigma \sum_{T\in A\cap\sigma} \left(\Pr[\xi \in T|\xi \in \sigma] - \Pr[\delta'_v \in T|\delta'_v \in \sigma]\right)\Pr[u \in H_\sigma]\Pr[w \in I_\sigma] \\
= \;& \sum_{H_\sigma} \Pr[u \in H_\sigma] \sum_{I_\sigma} \sum_{T\in(H_\sigma, I_\sigma)\cap A} \left(\Pr[\xi \in T|\xi \in \sigma] - \Pr[\delta'_v \in T|\delta'_v \in \sigma]\right)\Pr[w \in I_\sigma] \\
\leq \;& 2\cdot 2^{-c'm} \sum_{I_\sigma} \Pr[w \in I_\sigma] \leq 2^{-cm}
\end{aligned}$$

The last line follows from Corollary 3.2 for the strips σ of distance at most $\frac{mR}{40}$ from a coset of H intersecting L and from the fact that the probability that w is at distance greater than $\frac{mR}{40}$ is exponentially small in m.

Let $B = 2K^{1/c'}$ for some $c' > 1$. Let $C = (\frac{K}{2}, \frac{K}{2}, \ldots, \frac{K}{2})$. Let Q be the cube centered at C with sides of length $K - 2B$. Let Q' be the cube centered at C with sides of length $K - B$. Intuitively, for any strip σ intersecting Q, if we tile $\sigma \cap Q'$ with copies of $\mathcal{T}$, then for each tile T intersecting Q, $\Pr[\xi_{L,K,R} \in T]$ and $\Pr[\delta_v \in T]$ depend only on I_σ. By Lemma 3.9, the probability that either δ_v or $\xi_{L,K,R}$ belongs to a tile not intersecting Q is exponentially small in n. Restricting the argument above to those tiles whose intersection with Q is nonempty yields Part (1).

Proof of (2): The distance between the distributions η_K and δ_v is

$$\sup\{|\Pr(\eta_K \in A) - \Pr(\delta_v \in A)| + |\Pr(\eta_K \in B) - \Pr(\delta_v)| \text{ s.t. } A, B \in \Omega,\ A \cap B = \emptyset\}.$$

Let A be the set of tiles T for which $\Pr(\eta_K \in T) > \Pr(\delta_v \in T)$, let B be the set of tiles for which $\Pr(\delta_v \in T) \geq \Pr(\eta_K \in T)$ and let A' be those tiles whose intersection with Q is nonempty, where Q is as above. We analyze the first term.

$$
\begin{aligned}
& \sum_{T\in A'}(\frac{\text{volume}(T)}{\text{volume}(KU^n)} - (\Pr[\delta_v \in T|w_1 \text{ good}]\Pr[w_1 \text{ good}] + \Pr[\delta_v \in T|w_1 \text{ bad}]\Pr[w_1 \text{ bad}])) \\
\leq \; & \sum_{T\in A'}(\frac{\text{volume}(T)}{\text{volume}(KU^n)} - \Pr[\delta_v \in T|w_1 \text{ good}]\Pr[w_1 \text{ good}]) + \Pr[w_1 \text{ bad}] \\
\leq \; & \sum_{T\in A'}(\frac{\text{volume}(T)}{\text{volume}(KU^n)} - \Pr[\delta_v \in T|w_1 \text{ good}]\Pr[w_1 \text{ good}]) + 2^{-\frac{m}{4}\log\frac{m}{4}}
\end{aligned}
$$

by Corollary 3.1. If w is good, then since $u + \alpha v \in Q$ with probability at least $1 - 2^{-c'n}$ (Lemma 3.9), and with probability at least $1 - 2^{-c_4 n}$, $u + v$ is not in the same coset as u, it follows from Lemma 3.9 and Corollary 3.2 that the difference in Equation 2 is bounded by $2^{-c_5 n}$ for some $c_5 > 0$. Since the analysis of the second term in the expression for the distance between η_K and δ_v is analogous, we have the desired 2^{-cn} bound for some $c > 0$. ■

To finish the proof of the reduction we need to address the fact that (1) $\mathcal{A}$ and $\mathcal{B}$ can access only approximations to real numbers in a polynomial time; (2) $\mathcal{B}$ may not be able to sample $KU^n \cap L$ perfectly uniformly in polynomial time.

For (2), note that Lemma 3.10 holds even if w_0 (in the definition of δ_v) is not a copy of w but instead is obtained by choosing a point in $KU^n \cap L$ with a distribution whose distance from the uniform distribution on $KU^n \cap L$ is exponentially small. By Lemma 3.5, this can be done in polynomial time.

To address (1) we note that if two probability distributions on $\mathbb{R}^n$ are exponentially close to each other and we approximate each of these distributions to within $2^{-n^{e_0}}$ (a random vector chosen according to the given distribution is approximated by approximating each coordinate to within the given bound), then the respective approximations will themselves be exponentially close to each other. This completes the proof of Theorem 3.1.

Remark. The indistiguishability of $\xi_{L,K,R}$ from η_K yields a pseudo-random number generator.

4 Detailed Description of the Cryptosystem

4.1 The Public Key Cryptosystem Generator

To generate an instance of the cryptosystem, on input 1^n (the security parameter), we first choose d, M, K, R satisfying the requirements of Theorem 3.1 and generate a (d, M) lattice L represented by basis $Y = (y_1, \ldots, y_n)$. We then obtain a random basis B_L for L of length at most polynomial in n times d_L using Lemma 3.1. The public key is B_L together with K and R. The private key is Y. The hidden hyperplane H is the $n-1$ dimensional subspace defined by $y_1, \ldots, y_{n-1}$. Let $d' \geq d$ denote the distance of H to its nearest coset.

4.2 Encryption and Decryption

Let $B_L = (b_1, \ldots, b_n)$ and let B_L, K, R denote the public key. To encrypt a zero, randomly choose $\xi_{L,K,R}$ as described in Section 3.2. To encrypt a one, randomly choose η_K as described in Section 3.2.

Let u_H be a unit vector orthogonal to the subspace H, and let d_L be the distance of the consecutive hyperplanes. To decrypt the ciphertext z, the receiver computes the fractional part of $(u_H \cdot z)/d_L$. If it is within $\frac{nR}{d_L}$ of 0 or 1 then z is decrypted as 0, and as 1 otherwise.

By Theorem 3.1, if $\mathcal{A}$ is a probabilistic algorithm which distinguishes $\xi_{L,K,R}$ and η_K on X with a probability of at least $\frac{1}{2} + n^{-c_2}$, then, there exists a probabilistic algorithm $\mathcal{B}$, that using $\mathcal{A}$ as an oracle, finds $L^{(d,M)}$ on X with a probability of at least $1 - 2^{-n}$, in time n^{c_3}. It follows that the cryptosystem is secure if a random instance of the hidden hyperplane problem is hard.

5 Extension to General Lattices

We now describe how to remove the assumption that the (d, M) lattices are presented by a basis of length at most $n^c d$ for some $c > 0$. In Theorem 3.1 we showed that, given a distribution $\mathcal{L}$ on (d, M) lattices presented by bases of length at most a polynomial (in n) factor greater than d, $L \in_R \mathcal{L}$, a polynomial time algorithm distinguishing the two distributions $\xi_{L,K,R}$ and η_k could be used to effectively find the $n-1$ dimensional subspace $H = L^{(d,M)}$. In order to do this we sampled lattice points v from a small cube, used these v to generate distributions, and tested using, the distinguisher, if the v's were in H. Because the cube was small, it was intersected by only a polynomial number of hyperplanes, so in polynomial time we were able to find a basis for H.

The sampling used the fact that the length of the basis for L exceeded d by a factor of at most n^c for some $c > 0$. By eliminating this requirement we can no longer sample inside a small cube. We get around this problem by using the distinguisher to help us find random short vectors very close to H, and then then "growing" these into long vectors, still quite close to H. The growing takes place in stages; we use the distinguisher at every stage to recognize when a vector close to H has been found.

The long vectors are then used to find an approximation to H. If the approximation is sufficiently good then the unit vector orthogonal to the approximation will be very close to the unit vector u_H orthogonal to H. If the two vectors are sufficiently close then u_H can be found by rounding the unit vector orthogonal to the approximation.

Theorem 5.1 *There exist $c, c_4, c_5, c_6 > 0$ so that for all $c_1 > 0$, $c_2 > 0$ there exists $c_3 > 0$ and a probabilistic algorithm $\mathcal{B}$ (using an oracle) so that if n, d, M, K, R are positive integers satisfying the inequalities,*
(1) $\log d + \log M + \log K + \log R < n^{c_1}$
(2) $n^{c_6} M > d > n^{c_4} M$,
(3) $R > n^c M$,

(4) $K > 2^{c_5 n} d$,
and $\mathcal{L}$ *is a distribution on* (d, M)*-lattices in* $\mathbb{Z}^n$ *presented by vectors in a cube of size* $2^{n^{c_1}} d$, *and* $\mathcal{A}$ *is a probabilistic polynomial time algorithm which distinguishes* $\xi_{L,K,R}$ *and* η_K *on* $\mathcal{L}$ *with a probability of at least* $\frac{1}{2} + n^{-c_2}$, *then,* $\mathcal{B}$, *using* $\mathcal{A}$ *as an oracle, finds* $L^{(d,M)}$ *on* $\mathcal{L}$ *with a probability of at least* $1 - 2^{-n}$, *in time* n^{c_3}.

Remark 5.1 *Since* M *is just an upper bound on the length of* $L^{(d,M)}$ *the requirement that* $n^{c_6} M > d$ *does not restrict* $\mathcal{L}$.

Proof: Let w be a random variable whose value is $w = \sum_{i=1}^{m} t_i$, where for $1 \le i \le m$, $t_i \in_R S^n(n^3 M)$. Let u be the random variable defined by choosing uniformly a vector in $KU^n \cap L$. Let the random variable $\xi_{L,K,R}$ be defined by $\xi_{L,K,R} = u + w$. Let r be the random variable defined by choosing uniformly a vector in $KU^{(n)} \cap \mathbb{R}^n$ within distance z of a coset of $L^{(d,M)}$ intersecting L. Let $\xi'_{L,K,R,z}$ be the random variable $\xi'_{L,K,R,z} = r + w$.

Let $\alpha \in_R [0,1]$. For $v \in \mathbb{R}^n$, let δ_v be the random variable defined by $\delta_v = u + \alpha v + w$. In general, for $v \in \mathbb{R}^n$, let the perturbation $g(v)$ be the random variable v by $g(v) = v + w$.

The next lemma shows that a distinguisher $\mathcal{A}$ satisfying the conditions of the Theorem can be used to separate vectors close to H from those close to any coset $H_1 \neq H$ of H intersecting L.

Lemma 5.1 *For all* $c > 0$ *there exist* $c_1 > 0$ *and* $r_c \ge R/n^{c_1}$ *such that*

1. *if* v *is within distance* r_c *of* H, *then*

$$|\Pr[\mathcal{A}(\delta_v) = 1] - \Pr[\mathcal{A}(\xi_{L,K,R}) = 1]| < \frac{1}{n^c}$$

2. *if* v *is within distance* r_c *of a coset* $H_1 \neq H$ *in* L, *then*

$$|\Pr[\mathcal{A}(\delta_v) = 1] - \Pr[\mathcal{A}(\eta_K) = 1]| < \frac{1}{n^c}$$

Proof: Consider the distribution $\xi'_{L,K,R,0}$ obtained by perturbing vertices $v \in KU^{(n)}$ that are not in the lattice but that lie in the hyperplane H or one of its cosets intersecting L. By Lemma 3.10 this distribution is differs from $\xi_{L,K,R}$, the distribution induced by perturbing randomly chosen vertices $v \in KU^{(n)} \cap L$, by at most 2^{-cn} for some $c > 0$.
Proof of (1)

We calculate an upper bound on the difference between the distribution $\xi'_{L,K,R,0}$ and $\xi'_{L,K,R,z}$ as a function of z. For any $p \in \mathbb{R}^n$, the *height* of p denotes the distance of p to the nearest coset of H in L. By definition, the two distributions differ only in the distribution on the heights of the sampled points.

Let t be chosen uniformly at random from $S^n(1)$, and let t' be chosen uniformly at random from an n-dimensional ball of the same radius but centered at a point at height ζ. We will compare the difference in height between t and t'.

The probability density function determining the height of t is proportional to $(\sqrt{1-x^2})^{n-1}$. So there exists a constant $c' = c'(n)$ satisfying $\int_{-1}^{1} c'(1-x^2)^{\frac{n-1}{2}} dx = 1$. We bound c' as follows. In this range, $1 - \frac{1}{x^2} > 1 - \frac{1}{n}$, so $(1-\frac{1}{x^2})^{\frac{n-1}{2}} > (1-\frac{1}{n})^{\frac{n-1}{2}}$, which is approximately $e^{-\frac{1}{2}}$. So $\int_{-1}^{1}(1-x^2)^{\frac{n-1}{2}} dx \geq \int_0^{1/n}(1-x^2)^{\frac{n-1}{2}} \geq \frac{1}{n}e^{-\frac{1}{2}}$, and $c' \leq ne^{\frac{1}{2}}$.

We next bound the difference between the two intervals $\int_{-1}^{1}(1-x^2)^{\frac{n-1}{2}} dx$ $\int_{-1+\zeta}^{1+\zeta}(1-x^2)^{\frac{n-1}{2}} dx$.

$$\begin{aligned}
\frac{d}{dx}(R^2-x^2)^{\frac{n-1}{2}} &= \frac{n-1}{2}(R^2-x^2)^{\frac{n-3}{2}}(-2x) \\
&= (1-n)(R^2-x^2)^{\frac{n-3}{2}}x \\
\frac{d^2}{dx}(R^2-x^2)^{\frac{n-1}{2}} &= (1-n)[(R^2-x^2)^{\frac{n-3}{2}} + (n-3)(R^2-x^2)^{\frac{n-5}{2}}(-x^2)]
\end{aligned}$$

The second derivative is zero when $(R^2-x^2)^{\frac{n-3}{2}} = x^2(n-3)(R^2-x^2)^{\frac{n-5}{2}}$, which simplifies to $(R^2-x^2) = (n-3)x^2$, or $x = \frac{\pm R}{\sqrt{n-2}}$. Using this value for x we get that the slope of the curve $f(x) = (\sqrt{1-x^2})^{n-1}$ is bounded by $\frac{n-1}{\sqrt{n-2}}(1-\frac{1}{n-2})^{\frac{n-3}{2}}$, or approximately $\sqrt{n}e^{-\frac{1}{2}}$. So for any x the magnitude of the difference between $f(x)$ and $f(x+\zeta)$ is bounded by $\zeta\sqrt{n}e^{-\frac{1}{2}}$. Integrating this difference over $[-1, 1+\zeta]$ yields an amount bounded by $3\zeta\sqrt{n}e^{-\frac{1}{2}}$. To ensure a difference bounded by a given polynomial n^{-c_1}, we need only choose $\zeta = n^{-c_1-\frac{1}{2}}$. To scale to a ball of radius R we take $z = \zeta R$.

Intuitively, if v is within z of H then so is αv, and $\delta_v = u + \alpha v + w$ is essentially the same distribution as $\xi'_{L,K,R,z}$ and hence is close to $\xi_{L,K,R}$. By choosing $r_c = z = \frac{1}{c'}n^{-(c+\frac{1}{2})}R$ we ensure that

$$|\Pr[\mathcal{A}(\delta_v) = 1] - \Pr[\mathcal{A}(\xi_{L,K,R}) = 1]| < \frac{1}{n^c}.$$

Proof of (2)

By assumption, $\exists c^* > 0$ such that $|\Pr[\mathcal{A}(\xi_{L,K,R}) = 1] - \Pr[\mathcal{A}(\eta_K) = 1]| \geq \frac{1}{n^{c^*}}$. Without loss of generality, assume $\Pr[\mathcal{A}(\xi_{L,K,R}) = 1] < \Pr[\mathcal{A}(\eta_K) = 1]$. Let $\textsc{small}' = \Pr[\mathcal{A}(\xi_{L,K,R}) = 1]$, $\textsc{small} = \Pr[\mathcal{A}(\xi'_{L,K,R,0}) = 1]$, and $\textsc{large} = \Pr[\mathcal{A}(\eta_K) = 1]$. By Lemma 3.10, $\textsc{small}'$ is exponentially (in n) close to $\textsc{small}$.

Let the random variable $g(\eta_K)$ be defined as follows. Choose $p \in_R \eta_K$ Then $g(\eta_K) = p+w$, where w is the random variable defined above. $g(\eta_K)$ is uniformly distributed in KU^n, so $\Pr[\mathcal{A}(\eta_K) = 1] = \Pr[\mathcal{A}(g(\eta_K)) = 1]$.

Consider the random variable $\delta_v = u + \alpha v + w$. If v is within distance z of a coset $H' \neq H$ of H intersecting L, then, conditioned on αv not being within z of such a coset, the distance of $u + \alpha v$ to the nearest coset is uniformly distributed in $[z, \frac{d}{2}]$. If αv is within z of a coset, then $u + \alpha v + w$ has essentially the same distribution as $\xi'_{L,K,R,z}$: it is a perturbation on the point $u + \alpha v$.

Let us use the notation $\textsc{close}(p)$ to indicate that p is within z of the nearest coset of H intersecting L.

$$\Pr[\mathcal{A}(\eta_K) = 1] = \Pr[\mathcal{A}(g(\eta_K)) = 1]$$

$$
\begin{aligned}
&= \Pr[\mathcal{A}(g(\eta_K)) = 1 | \text{CLOSE}(p)] \; \Pr[\text{CLOSE}(p)] \\
&\quad + \Pr[\mathcal{A}(g(\eta_K)) = 1 | \neg\text{CLOSE}(p)] \; (1 - \Pr[\text{CLOSE}(p)])
\end{aligned}
$$

In other words, $\text{LARGE} = \text{SMALL} \cdot \frac{2z}{d} + \rho \cdot (1 - \frac{2z}{d})$, where ρ is the probability that $\mathcal{A}(g(\eta_K)) = 1$, given that p is not within z of a coset of H intersecting L. By appropriate choice of z we can ensure that $|\text{LARGE} - \rho < n^{-c}|$, completing the proof of the lemma.

By assumption, $\text{LARGE} - \text{SMALL} \geq n^{-c_2}$, so $\text{LARGE} - n^{c_2} \geq \text{SMALL}$. We also have $\text{LARGE} = \text{SMALL} \cdot \frac{2z}{d} + \rho \cdot (1 - \frac{2z}{d})$. Simple algebraic manipulation yields

$$
\begin{aligned}
\text{LARGE} - \rho(1 - \frac{2z}{d}) &= \text{SMALL}\frac{2z}{d} \\
&\leq (\text{LARGE} - n^{c_2})\frac{2z}{d} \\
|\text{LARGE} - \rho| &\leq n^{-c_2}\frac{2z/d}{1 - 2z/d}
\end{aligned}
$$

To ensure that $|\text{LARGE} - \rho| < n^{-c}$ it suffices to choose $z < n^{-c+c_2}\frac{d}{4}$. Part (2) follows from the fact that $R < d$. ■

Let σ denote the set of all points in $\mathbb{R}^n$ of distance at most d from H. The following iterative process finds long vectors close to H.

Growing Long Vectors

Each iteration has a *starting point* s which for the first iteration is the origin, and in general will always be within distance $2d$ of H. Let $S(2\sqrt{n}d, s)$ be a ball of radius $2\sqrt{n}d$ around the starting point s. The goal is to find a point v in $\sigma \cap S(2\sqrt{n}d, s)$ that is farther from the origin than s and still inside σ. Then $2v$ becomes the new starting point, and the process continues. Occasionally the procedure may err; this is eventually detected and the computation is backed up to an earlier starting point and repeated with different random choices.

In Section 3.2 we used the distinguisher to test points $v \in L$ to see if they are outside of H. Specifically, v was tested by sampling the distribution δ_v and testing $\mathcal{A}$ on the samples. We will use the same test here, this time to distinguish points near H from points outside of σ. Specifically, we have a way of choosing random points v within distance $2d$ of H and testing them such that: (1) if $v \notin \sigma$ then with high probability this is detected; (2) if v is "very close" to H then with high probability v is recognized as being in σ; and (3) the probability that we find a $v \in \sigma$ that is not falsely detected as being outside of σ is polynomial in n^{-1}.

Our goal is to construct an approximation $\hat{H}$ to H by finding $n-1$ mutually orthogonal long lattice vectors $v_1, \ldots, v_{n-1}$, say, of length at least ℓ, all at distance less than d of H. Once we have found $v_1, \ldots v_{i-1}$, we search for v_i in the $n-i+1$ dimensional subspace V^{n-i+1} of $\mathbb{R}^n$ orthogonal to $v_1, \ldots, v_{i-1}$, such that v_i is close to $H \cap V^{n-i+1}$.

We now describe the general step of searching for the next starting point in the construction of v_i.

Finding the Next Starting Point

Let $S^{n-i+1}_{(v_1,\ldots,v_{i-1})}(2\sqrt{n}d)$ denote the $n-i+1$ dimensional ball in the $n-i+1$ dimensional subspace orthogonal to $v_1,\ldots,v_{i-1}$ of radius $2\sqrt{n}d$ centered at the origin. Choose a random $v' \in S^{n-i}_{(v_1,\ldots v_{i-1})}(2\sqrt{n}d)$ such that $||s+v'|| > ||s||$. Test if $s+v'$ is outside of σ, by testing each of $s+v'$, $s+\frac{n^{c_1}-1}{n^{c_1}}v'$, $s+\frac{n^{c_1}-2}{n^{c_1}}v'$, $\ldots s+\frac{1}{n^{c_1}}v'$ to see if any is near a coset $H_1 \neq H$ in L. (For any vector u this test is accomplished by sampling from δ_u (cf. Lemma 5.1).) If any multiple of v' tests positive, then v' is discarded and the procedure is repeated for a new random v'. If no test is positive and $\|s+v'\| \geq \ell$, then we set $v_i = s+v'$. If no test is positive but $\|s+v'\| < \ell$, then we set the next starting point to $2(s+v')$.

Lemma 5.2 *There exist $c_3, c_4, c_5 > 0$ such that with probability $1-n^{-c_4}$, if the procedure described above is run from a starting point s within distance $2d$ of H, then within n^{c_5} iterations the procedure produces an output that, with probability $1-n^{c_3}$, is in σ.*

Proof: Recall that $n^{c_6}M > d$ for some $c_6 > 0$, and that $R = n^3 M$. So for some $c_1 > 0$, $d \leq n^{c_1}R$. Let $\alpha = \frac{1}{c'}n^{-c+\frac{1}{2}+c_1}$. Let n^{-c_2} be the assumed distinguishing advantage of $\mathcal{A}$.

Let e_0 be the probability that the test errs on inputs within distance αd of $H \cap V^{n-i+1}$. Let e_1 be the probability that the test errs on inputs within distance $n^{-c+c_2}\frac{d}{4}$ of a coset of $(H' \cap V^{n-i+1}) \neq (H \cap V^{n-i+1})$ of $H \cap V^{n-i+1}$ in L.

If $s+v'$ is within distance αd of $H \cap V^{n-i+1}$ then none of the multiples tested yields the uniform distribution, so, with probability $1-n^{c_1}e_0$, v will not be discarded. If $s+v'$ is outside σ then at least one of the multiples is within distance $\frac{d}{4}n^{-c+c_2}$ of $H \cap V^{n-i+1}$ (this uses the fact that $||v'|| \leq 2\sqrt{n}d$), so, with probability at least $1-e_1$, v' will be discarded.

It remains only to show that with sufficiently high probability for a random $v' \in S^{n-i}_{(v_1,\ldots,v_{i-1})}(2\sqrt{n}d)$, the vector $s+v'$ will be within αd of H. (We need to be sure there is sufficient probability of finding a vector this close to H because it is only on these vectors that we are "guaranteed" that the distinguisher will recognize an encryption of 0. The actual vector produced may be anywhere in σ.)

Let $Y = 2\sqrt{n}d$ and let $m \geq 2 \in \mathbb{Z}$. Let G be an $m-1$ dimensional hyperplane passing through the origin. Consider the m-dimensional ball of radius Y centered at a point p of distance at most $2d$ from G. We bound the ratio of the volume of the ball to the volume of the strip of points in $\mathbb{R}^m$ parallel to G and at distance $2d \leq y \leq 3d$ from an $m-1$ dimensional hyperplane parallel to G passing through p. The probability of choosing a point within αd of G is at least $\alpha d/d$ times this ratio.

The ratio is given by

$$\frac{\int_{2d/Y}^{3d/Y}(\sqrt{1-x^2})^{m-1}dx}{\int_0^1(\sqrt{1-x^2})^{m-1}dx}$$

The numerator is bounded below by $\frac{1}{2\sqrt{n}}(1-\frac{9}{4n})^{\frac{m-1}{2}} \geq \frac{1}{2\sqrt{n}}e^{-\frac{9}{2}}$ because $m \leq n$, while the denominator can be shown to be bounded above by $c''\sqrt{m}$ for a constant $c'' > 0$ independent of m. Multiplying by α yields a probability polynomial in $\frac{1}{n}$. ∎

Corrolary 5.1 *There exists c_5 such that and for all $c_6 \geq 0$ there is procedure that, using $\mathcal{A}$ as an oracle, with probability at least $1 - n^{-c_5}$ generates a vector v within distance d of H and having length at least $2^{n^{c_7}}$ in time polynomial in n^{c_6}.*

Using Corollary 5.1, we can find $n-1$ mutually orthogonal long vectors $u_1, \ldots, u_{n-1}$ close to H. We let $\hat{H}$ denote the approximation to H defined by $\{u_1, \ldots, u_{n-1}\}$.

We measure the quality of the approximation by finding the distance between the unit vectors orthogonal, respectively, to H and $\hat{H}$. Given $n-1$ vectors $u_1, \ldots u_{n-1} \in \mathbb{R}^n$, define a generalization of the cross product $\bigotimes(u_1, \ldots, u_{n-1})$ to be the vector in $\mathbb{R}^n$ whose ith coordinate is the determinant of the minor M_{1i} of the $n \times n$ matrix A with rows $e, u_1, \ldots u_{n-1}$. The key point is that for any $v \in \mathbb{R}^n$,

$$v \cdot \bigotimes(u_1, \ldots, u_{n-1}) = \begin{vmatrix} v_1 & v_2 & \ldots & v_n \\ u_{11} & u_{12} & \ldots & u_{1n} \\ \vdots & & & \vdots \\ u_{n-1,1} & u_{n-1,2} & \ldots & u_{n-1,n} \end{vmatrix}.$$

In particular, for $1 \leq i \leq n-1$, $u_i \cdot \bigotimes(u_1, \ldots, u_{n-1}) = 0$. Let x be a unit vector in the direction of $\bigotimes(u_1, \ldots, u_{n-1})$. Then

$$\begin{aligned} x \cdot \bigotimes(u_1, \ldots, u_{n-1}) &= \|x\| \, \|\bigotimes(u_1, \ldots, u_{n-1})\| \cos(0) \\ &= \|\bigotimes(u_1, \ldots, u_{n-1})\|. \end{aligned}$$

But $x \cdot \bigotimes(u_1, \ldots, u_{n-1}) = \det(x, u_1, \ldots u_{n-1})$ which is the volume of the parallelepiped $\mathcal{P}(x, u_1, \ldots u_{n-1})$, which, since $\|x\| = 1$, is the volume of the parallelepiped $\mathcal{P}(u_1, \ldots, u_{n-1})$. So, since $\|x\| = 1$, $\|\bigotimes(u_1, \ldots u_{n-1})\|$ equals the volume of the parallelepiped $\mathcal{P}(u_1, \ldots, u_{n-1})$. Finally, $\bigotimes(u_1, \ldots u_{n-1})$ has positive orientiation: $\det(\bigotimes(u_1, \ldots u_{n-1}), u_1, \ldots u_{n-1}) = \bigotimes(u_1, \ldots, u_{n-1}) \cdot \bigotimes(u_1, \ldots, u_{n-1}) \geq 0$, so the cross product has positive orientation unless it is zero.

Let us assume that we have a basis for $\mathbb{R}^n$ in which the nth basis vector is u_H, a unit vector orthogonal to H. For $1 \leq i \leq n$, our ith basis vector can be written as $u_i = (u_{i1}, \ldots, u_{i,n-1}, \epsilon_i)$, where by construction, each $|\epsilon_i| < d$. By appropriate choice of ℓ we can arrange that $|\epsilon_i|$ is small relative to $\|u_i\|$. Let x be the unit vector in the direction of $\bigotimes(u_1, \ldots, u_{n-1})$. Then the distance of x to u_H is given by $\sqrt{1 - x_n^2}$. Our goal is to show that $|x_n|$ is very close to 1.

Let V be the volume of the parallelepiped with sides $(x, u_1, \ldots, u_{n-1})$. By definition,

$$\begin{aligned} x \cdot \bigotimes(u_1, \ldots, u_{n-1}) &= \begin{vmatrix} x_1 & x_2 & \ldots & & x_n \\ u_{11} & u_{12} & \ldots & u_{1,n-1} & \epsilon_1 \\ \vdots & & & \vdots & \\ u_{n-1,1} & u_{n-1,2} & \ldots & u_{n-1,n-1} & \epsilon_{n-1} \end{vmatrix} \\ &= x_1 \det(M_1) - \ldots + (-1)^{n+1} x_n \det(M_n) \end{aligned} \tag{2}$$

where M_i is the $(1,i)$ minor of the matrix (expanding along the first row). Thus, $|V| = \|u_1\| \ldots \|u_{n-1}\| \cdot \|x\| = \|u_1\| \ldots \|u_{n-1}\|$. Let M_i^* denote the (i,n) minor of the matrix (expanding along the nth column). Then

$$\begin{aligned} & x_1 \det(M_1) - \ldots + (-1)^{n+1} x_n \det(M_n) \\ = \; & (-1)^{n+1} x_n \det(M_n) + (-1)^{n+2} \epsilon_1 \det(M_1^*) + \ldots + \epsilon_{n-1} \det(M_{n-1}^*). \end{aligned}$$

For $1 \leq i \leq n-1$, let $u_i^* = (u_{i1}, \ldots, u_{i,n-1})$. Since $\|x\| = 1$, $\det(M_i^*)$ is bounded by the $\mathcal{U}_i$, the volume of the parallelepiped with sides $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_{n-1}$. Let $V^* = \text{volume}(\mathcal{P}(u_1^*, \ldots, u_{n-1}^*))$, let $\epsilon = \max_{i=1}^{n-1} \epsilon_i$ and $\alpha = \min_{i=1}^{n-1} \|u_i^*\|$. Then

$$\begin{aligned} |V| & \leq |x_n V^*| + \sum_{i=1}^{n-1} |\epsilon_i \mathcal{U}_i| \\ & \leq |x_n V^*| + n |\frac{\epsilon V^*}{\alpha}| \\ \frac{|V|}{|V^*|} & \leq |x_n| + |\frac{n\epsilon}{\alpha}|. \end{aligned} \tag{3}$$

But $1 \leq \frac{|V|}{|V^*|}$, so $|x_n| \geq 1 - |\frac{n\epsilon}{\alpha}|$. Since $\epsilon \leq d$ and $\alpha \geq \frac{\ell}{2}$, we can make x_n as close to 1 as desired by appropriate choice of ℓ. In particular, since $u_H = (0, 0, \ldots, 0, 1)$, the distance from x to the unit vector orthogonal to H can be made as small as desired.

Lemma 5.3 *Assume that $b_1, ..., b_n$ is a basis of the lattice $L \subseteq \mathbb{R}^n$, $b'_1, ..., b'_n$ is its dual basis, $\|b'_i\| \leq N$ for $i = 1, ..., n$, $v \in L$, $u = \sum_{i=1}^n \beta_i b_i \in \mathbb{R}^n$, $\|u - v\| < \frac{1}{2N}$ and α_i is the closest integer to β_i for $i = 1, ..., n$. Then $\sum_{i=1}^n \alpha_i b_i = v$.*

Proof: $\beta_i = b'_i \cdot u$ (inner product) for $i = 1, ..., n$. If $v = \sum_{i=1}^n \gamma_i b_i$ then $\gamma_i = b'_i \cdot v$. It is enough to show that $|\beta_i - \gamma_i| < \frac{1}{2}$. $|\beta_i - \gamma_i| = |b'_i \cdot v - b'_i \cdot u| = |b'_i(v - u)| < N \frac{1}{2N} = \frac{1}{2}$. ■

We want to apply Lemma 5.3 to L^*, where $L \in_R \mathcal{L}$. Note that $L = (L^*)^*$ has a basis of length at most $N = 2nd$. $u_H \in L^*$ and x is close to u_H. The lemma says that if $\|x - u_H\| < \frac{1}{2N}$ then if we write x as a linear combination of the basis vectors for L^* and round the coefficients of these basis vectors to the nearest integers, we will obtain u_H. In order to ensure that $\|x - u_H\| \leq \frac{1}{2N}$ it is necessary and sufficient that $\sqrt{1 - x_n^2} < \frac{1}{2N}$. Since $|x_n| \geq 1 - |\frac{n\epsilon}{\alpha}| \geq 1 - |\frac{2dn}{\ell}|$, choosing $\ell > 16dN^2n$ suffices.

This completes the proof of Theorem 5.1. ■

Appendix 2
A Public-key Cryptosystem with Worst-case/Average-case Equivalence

Miklos Ajtai and Cynthia Dwork

We present a probabilistic public key cryptosystem which is secure unless the following worst-case lattice problem can be solved in polynomial time: "Find the shortest nonzero vector in an n dimensional lattice L where the shortest vector v is unique in the sense that any other vector whose length is at most $n^c\|v\|$ is parallel to v."

1 Introduction

In [3] we presented a public key cryptosystem generator whose security was based on the unique shortest vector problem:

> Find the shortest nonzero vector in an n dimensional lattice L where the shortest vector v is unique in the sense that any other vector whose length is at most $n^c\|v\|$ is parallel to v

in the sense that from each instance of the problem it was possible to create a public key cryptosystem so that (without the corresponding private key), distinguishing encryptions of zeros from encryptions from ones, when the encryptions are performed using this particular public key is just as difficult as to solve the given instance of the lattice problem. In this paper we give another cryptosystem generator so that the following holds: if a random instance of the system can be broken, that is, the probability that an encryption of a zero can be distinguished from an encryption of a one (without the private key) in polynomial time with a probability of at least n^{-c_1} for some absoloute constant $c_1 > 0$, then the worst-case unique shortest vector problem has probabilistic polynomial time solution. The unique shortest vector problem is one of the three problems listed in [2]. There, a random method is given to generate hard instances of a lattice problem so that if it has a polynomial time solution then all of the three worst-case problem (including the unique-shortest vector problem) has a solution. For further information about lattice problems and public cryptosystems we refer the reader to the introduction of [3].

2 The Main Theorem

The definitions and theorems of this paper will use three constants, $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$. We assume that $\mathcal{D}_1 = 3, \mathcal{D}_2 = 8, \mathcal{D}_3 = 3$. (We do not write this actual values in the definitions and statements since this way the results are more easily adaptable for other possible values of these constants. In a similar way $\mathcal{K}(n)$ will denote function $2^{n \log n}$. We made no attempt to choose these constants and the function in an optimal way in any sense.

Definitions. Let n be a fixed positive integer. Most of the definitions depend on n but the notation will not always show this dependence explicitly.

1. Let $\mathcal{Q}$ be the n-dimensional cube $\mathcal{K}U^{(n)}$ where $U^{(n)}$ is the n-dimensional unit cube. $\mathcal{U}'$ will be a random variable which takes its values with uniform distribution on $\mathcal{Q}$.

2. $\mathrm{pert}(R, m)$ is a random variable whose each value is the sum of m vectors taken independently and with uniform distribution from the n dimensional ball with radius R around 0.

3. Suppose that $u \in \mathbb{R}^n$, $0 < \|u\| \leq 1$, $R > 0$ and m is a positive integer. We define the random variable $\mathcal{H}'(u, R, m)$ in the following way:

First let X be the set of all $x \in \mathcal{Q}$ so that $x \cdot u$ is an integer. X consists of subsets of a finite number of $n-1$-dimensional hyperplanes, so the $n-1$ dimensional volume defined on these hyperplanes induces a probability measure on X. We take a random point y on X. Independently we also take a value z of the random variable $\mathrm{pert}(R, m)$. The value of $\mathcal{H}'(u, R, m)$ is $y + z$.

4. If $y \in \mathbb{R}$ and $\alpha > 0$ then let $\mathrm{round}_\alpha(y)$ be $i\alpha$, where i is the largest integer with $i\alpha \leq y$.

If $x = \langle x_1, ..., x_n \rangle \in \mathbb{R}^n$ then $\mathrm{round}_\alpha(x) = \langle \mathrm{round}_\alpha(x_1), ..., \mathrm{round}_\alpha(x_n) \rangle$.

5. Let $\mathit{unif} = \mathrm{round}_{2^{-n}}(\mathcal{U}')$ and $\mathcal{H} = \mathrm{round}_{2^{-n}}(\mathcal{H}')$

6. If $a_1, ..., a_n \in \mathbb{R}^n$, then $\mathrm{width}(a_1, ..., a_n)$ will be the width of the parallelepiped defined by the vectors $a_1, ..., a_n$, that is, the maximum of the distances between the point a_i and the subspace generated by $\{a_j | j \neq i\}$, for $i = 1, ..., n$.

7. If $k > 0$ is a fixed positive integer then we will call the following problem the n^k-unique shortest vector problem:

Find the shortest nonzero vector in an n dimensional lattice L where the shortest vector v is unique in the sense that any other vector whose length is at most $n^k \|v\|$ is parallel to v.

8. When we speak about an instance of the n^k-unique shortest vector problem we will assume that $L \subseteq \mathbb{Z}^n$ and L is represented by a basis. The sum of the logarithms of the absolute values of all of the nonzero components of the basis vectors will be called the size of the instance.

We describe a public cryptosystem, and prove that it is secure unless the $n^{\mathcal{D}_2}$-unique shortest vector problem has a polynomial time probabilistic solution.

We describe first a conceptually simple system, which also makes our proofs easier to understand. We will point out later that with some modification we can make it more efficient. We assume that a positive integer n is fixed and published. (Alternately n can be part of the public key.) The public and private keys are chosen in the following way.

(1) Choose a random vector u with uniform distribution on the set $\{x \in \mathbb{R}^n | \|x\| \leq 1\}$.

(2) The private key is u.

(3) Generate $m = n^{\mathcal{D}_3}$ independent values $v_1, ..., v_m$ of the random variable $\mathcal{H}_{u, n^{-\mathcal{D}_1}, n}$,

(4) $v_1, ..., v_m$ is the public key

For the encryption of a message, the sender will need the smallest integer i_0 so that width$(v_{i_0+1}, ..., v_{i_0+n})$ is at least $n^{-2}\mathcal{K}$. We will show later, that with a probability exponentially close to 1, $i_0 < n^2$. Since the value of i_0 does not depend on the message, we may consider i_0 also as a part of the public key. Let $w_1 = v_{i_0+1}, ..., w_n = v_{i_0+n}$. A $0, 1$ bit z will be encrypted in the following way:

if $z = 0$ then

(1) choose m random $0, 1$ values $\delta_1, ..., \delta_m$, independently and with probabilities $\frac{1}{2}, \frac{1}{2}$,

(2) compute the vector $x = \sum_{j=1}^{m} \delta_i v_i$,

(3) reduce the vector x modulo $w_1, ..., w_n$, into the praralellepiped $\mathcal{P}^{-}(w_1, ..., w_n) = \{\sum_{j=1}^{n} \lambda_j w_j | 0 \leq \lambda_j < 1, j = 1, ..., n\}$ that is, find the unique vector x' in $\mathcal{P}^{-}(w_1, ..., w_n) =$ so that $x - x'$ is an integer linear combination of the vectors $w_1, ..., w_n$.

(4) x' is the encrypted message.

If $z = 1$ then we get the encrypted message in the following way. Let $2^{-n}\mathbb{Z}^n$ be the set of all vectors of the form $2^{-n}b$, where $b \in \mathbb{Z}^n$.

(5) the encrytped message x'' is a random vector chosen with uniform distribution from the set $\mathcal{P}^{-}(w_1, ..., w_n) \cap 2^{-n}\mathbb{Z}^n$.

The decryption of the message. Assume that z is the encrypted message. Knowing the vector u we can decrypt the message in the following way. Let $u \cdot z = i + \theta$ where $i \in \mathbb{Z}$ and $-\frac{1}{2} \leq \theta < \frac{1}{2}$, where $u \cdot z$ is the inner product of the n dimensional vectors u and z. If $|\theta| < \frac{1}{n}$ then z is decrypted as a 0 otherwise as a 1. (Our definitions imply that if the bit 0 is sent then it is always decrypted correctly, while the bit 1 is decrypted correctly with a probability of at least $1 - \frac{1}{n}$).

We will show that our assumption about the hardness of the $n^{\mathcal{D}_2}$-unique shortest vector problem implies that the distribution of x' and x'' cannot be disntiguished by polynomial time computation.

x' and x'' as given in the definition of the encrypted message depend only on the vectors $v_1, ..., v_m$. For later use we define two random variables whose values can be computed in the same way as x' and x''.

Definition. Assume that $v_1, ..., v_m \in \mathbb{R}^n$, where $m = n^{\mathcal{D}_3}$. We define two random variables $\mathcal{S}_{v_1,...,v_m}$ and $\mathcal{E}_{v_1,...,v_m}$ in the following way. Let i_0 be the smallest integer so that width$(v_{i_0+1}, ..., v_{i_0+n})$ is at least $n^{-2}\mathcal{K}$ and let $w_1 = v_{i_0+1}, ..., w_n = v_{i_0+n}$. (If there is no i_0 with this porperty then the value of both random variables is 0.) We define x' and x'' in the same way as in the description of the protocol. x' will be a value of $\mathcal{S}_{v_1,...,v_m}$, x'' will be a value of $\mathcal{E}_{v_1,...,v_m}$.

Definition. Assume that ξ_0, ξ_1 are random variables whose each value is represented by a $0, 1$-sequence of length k, S is a $0, 1$ sequence of length s and $\mathcal{A}$ is a probabilistic circuit with $2kt + s$ input nodes and $\delta > 0$. We will say that $\mathcal{A}$ distinguishes the random variables ξ_0 and ξ_1 over S with a bias of at least δ if the following holds.

The input nodes of $\mathcal{A}$ are partitioned into three subsets Y_0, Y_1, Y_2 so that $|Y_0| = |Y_1| = kt$. Y_0 and Y_1 are further partitioned into blocks of sizes k. Suppose that we give t mutually indepent values of both ξ_0 and ξ_1 and the sequence S as inputs to $\mathcal{A}$ in the following way.

First we pick an $i \in \{0,1\}$. The values of ξ_i will be given on the blocks of Y_0 and the values of ξ_{i-1} on the blocks of Y_1. S will be given as input on Y_2. Then for each fixed choice of $i = 0,1$, the probability that the output of $\mathcal{A}$ is i is at least $\frac{1}{2} + \delta$, where the probability is taken for both the randomization of ξ_1, ξ_2 and the probabilistic steps in $\mathcal{A}$.

The following theorem says that $\mathcal{S}_{v_1,\ldots,v_m}$ and $\mathcal{E}_{v_1,\ldots,v_m}$ will be indistinguishable over $v_1, \ldots, v_m$ in this sense by any polynomial size circuit $\mathcal{A}$ which does not depend on the choice of u, provided that the $n^{\mathcal{D}_2}$-unique shortest vector problem has no polynomial time probabilistic solution. We assume that $v_1, \ldots, v_m$ is represented by a $0,1$ sequence. The fact that both these vectors and the values of ξ_0 and ξ_1 are elements of $2^{-n}\mathbb{Z}$ with a known upper bound on their components, provides a natural way for representing them by $0,1$-sequences.

Theorem 2.1 *For all $c_1, c_2, c_3, c_4 > 0$ there exists a c_5 and a probabilistic algorithm $\mathcal{B}$ (using an oracle) so that for all sufficiently large n, condition (1) implies condition (2), where*

(1) $\mathcal{A}$ is a probabilistic circuit of size n^{c_1} so that if $u, v_1, \ldots, v_m$ are picked at random as described in the protocol for generating the public and private keys, then with a probability of at least n^{-c_2} the following holds:

$\mathcal{A}$ distinguishes the random variables $\mathcal{S}_{v_1,\ldots,v_m}$ and $\mathcal{E}_{v_1,\ldots,v_m}$, over $v_1, \ldots, v_m$ with a bias of at least n^{-c_3}.

(2) $\mathcal{B}$, using $\mathcal{A}$ as an oracle, can solve any instance of size at most n^{c_4} of the $n^{\mathcal{D}_2}$-unique shortest vector problem in time n^{c_5} and with a probability at least $1 - 2^{-n}$.

Remarks. 1. As we have already indicated in the introduction,there are ways to make the cryptosystem more efficient. One possibility is that instead of choosing the vectors $w_1, \ldots, w_m$ from the set $v_1, \ldots, v_m$, we may randomize them separately, making sure that they are almost orthogonal, and so width$(w_1, \ldots, w_n)$ will be automatically large. E.g. we may pick w_i from the cube $\mathcal{K}e_i + (\mathcal{K})^{\frac{1}{2}} U^{(n)}$. The proof remains essentially the same, although it causes some additional complication that the vectors $v_1, \ldots, v_m, w_1, \ldots, w_n$ are now choosen (independently and) with $n+1$ different distributions.

2. Lemma 3.3 shows that if the worst-case n^c-unique short vector problem has no polynomial time solution then in a large cube the uniform distribution is computationally indistigushable from the following distribution:

we fix a random vector u in the unit ball, than take random points from the large cubes on the hyperplanes where the inner products of the vectors with u is an integer, and then perturb these points slightly. (Each perturbed point is a value of the random variable.)

Using the fact that the distribution of this random variable is computationally indistingushable from the uniform distribution we may construct a pseudo random number generator in the same way as it is done in [3].

3 Proof of the Theorem

Lemma 3.1 *For all c_1, c_2 there exists c_3, c_4 and a probabilistic algorithm $\mathcal{B}$ (using an oracle) so that for all sufficiently large n, condition (1) implies condition (2), where*

(1) $\mathcal{A}$ is a probabilistic circuit so that $u, v_1, ..., v_m$ are picked at random as described in the protocol for generating the public and private keys, then with a probability of at least n^{-c_1} the following holds:

$\mathcal{A}$ distinguishes the random variables $\mathcal{S}_{v_1,...,v_m}$ and $\mathcal{E}_{v_1,...,v_m}$, over $v_1, ..., v_m$ with a bias of at least n^{-c_2}.

(2) Suppose that we pick the random vector u as described in the protocol for generating the public and private keys. Then with a probability of at least n^{-c_3} the following holds: $\mathcal{B}$, using $\mathcal{A}$ as an oracle, distinguishes the random variables unif and $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ with a bias of at least $n^{-\frac{1}{4}}$, in time n^{c_5}.

Proof of Lemma 3.1.

We may assume that condition (1) of the lemma holds even if in its conclusion we require that $\mathcal{A}$ distinguishes x' and x'' with a bias of at least $\frac{1}{2}-n^{-2}$. Indeed suppose that $u, v_1, ..., v_m$ is fixed with the property that $\mathcal{A}$ distinguishes $\mathcal{S}_{v_1,...,v_m}$ and $\mathcal{E}_{v_1,...,v_m}$ with a bias of at least n^{-c_2}. $\mathcal{B}$ produces a long independent sequence of values of both random variables (n^{c_2+1} times longer, than required by $\mathcal{A}$) and applies $\mathcal{A}$ n^{c_2+1} times, then takes the majority of the decisions. The bias of the decision will be exponentially close to $\frac{1}{2}$, certainly greater than $\frac{1}{2}-n^{-2}$. For the sake of notational simplicity we assume that already the original $\mathcal{A}$ has this property.

Let X be the set of sequences $v_1, ..., v_m$ with the property that $\mathcal{A}$ distinguishes the random variables $\mathcal{S}_{v_1,...,v_m}$, $\mathcal{E}_{v_1,...,v_m}$ with a bias of at least $\frac{1}{2}-n^{-2}$. Since $\mathcal{S}_{v_1,...,v_m}$, $\mathcal{E}_{v_1,...,v_m}$ can be generated in polynomial time, there is a polynomial time probabilistic algorithm which, for any fixed values $v_1, ..., v_m$, approximates the bias of $\mathcal{A}$ with a polynomially small error, using only a polynomial number of applications of $\mathcal{A}$. Therefore there is a set $Y \supseteq X$ so that

(a) $y \in Y$ can be decided in polynomial time, with a probability exponentially close ot one.

(b) for each $\langle v_1, ..., v_m \rangle \in Y$, $\mathcal{A}$ distinguishes $\mathcal{S}_{v_1,...,v_m}$, $\mathcal{E}_{v_1,...,v_m}$ with a bias of at least $\frac{1}{2}-n^{-1}$.

Now we give a definition for the algorithm $\mathcal{B}$. $\mathcal{B}$ gets mt, $t = n^{2c_1}$ values of a random variable ξ and it tries to decide whether it is *unif* or $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ in the following way: $\mathcal{B}$ partitions the values into blocks of size m. For each fixed block B it computes two $0, 1$ bits $f(B)$ and $g(B)$. Assume that the values in B are $b_1, ..., b_m$. If $\langle b_1, ..., b_m \rangle \notin Y$ then $f(B) = g(B) = 0$. Suppose $\langle b_1, ..., b_n \rangle \in Y$. $\mathcal{B}$ produces as many independent values of the random variables $\mathcal{S}_{b_1,...,b_m}$, $\mathcal{E}_{b_1,...,b_m}$ as needed for the input of $\mathcal{A}$. $\mathcal{B}$ gives the values of $\mathcal{S}$ and $\mathcal{E}$ to $\mathcal{A}$ as an input in a random order. If $\mathcal{A}$ identifies $\mathcal{S}$ and $\mathcal{E}$ correctly then $g(B) = 1$ otherwise $g(B) = 0$. $f(B) = 1$ in both cases. Finally let $f_0 = \sum f(B)$, $g_0 = \sum g(B)$, where we take the sums for all t blocks. If $f_0 > \frac{1}{2}n^{-c_1}t$ and $g_0 > \frac{3}{4}f_0$ then $\mathcal{B}$ decides that $\xi = \mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ otherwise it decides that $\xi = unif$. This completes the definition of $\mathcal{B}$.

For any fixed u let B_u be the following event: if we randomize $v_1, ..., v_m$ as described in the protocol, then with a probability of at least n^{-2c_1}, we have that, "$\mathcal{A}$ distinguishes the random variables $\mathcal{S}$ and $\mathcal{E}$, over $v_1, ..., v_m$ with a bias of at least $\frac{1}{2}-n^{-1}$".

Condition (1) of the lemma implies that if we randomize only u then $P(B_u) \geq n^{-2c_1}$. Since $n^{-2c_1} > n^{-c_3}$, it is sufficient to show that if B_u holds then $\mathcal{B}$ distinguishes $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ and *unif* with a bias of at least $\frac{1}{4}$.

We will show that if $\xi = \mathcal{H}_{u,n^{-\mathcal{D}_1},n}$, then for each fixed block B the following holds, where the probabilities are taken both for the randomization of the elements $b_1, ..., b_m$ and the random steps of $\mathcal{B}$ and $\mathcal{A}$:

(c) $P(f(B) = 1) \geq n^{-c_1}$, $P(g(B) = 1 | f(B) = 1) \geq 1 - n^{-1}$.

Since the events for different blocks are independent These inequalities imply that with a probability exponentialy close to one, we have $f_0 > \frac{1}{2} n^{-c_1} t$ and $g_0 > \frac{3}{4} f_0$.

For $\xi = unif$ we will show that

(d) $P(g(B) = 1 | f(B) = 1) < \frac{1}{2} + \frac{1}{8}$.

This implies that with a probability exponentially close to one either $f_0 < \frac{1}{2} n^{-c_1} t$ or $g_0 < \frac{3}{4} f_0$.

Therefore it is enough to show that the inequalities (c) and (d) hold for the appropriate choice of ξ.

Assume first that $\xi = \mathcal{H}_{u,n^{-\mathcal{D}_1},n}$. Since $Y \supseteq X$ and $P(X) \geq n^{-c_1}$, we have $P(Y) \geq n^{-c_1}$. Since $\langle b_1, ..., b_m \rangle \in Y$ implies $f(B) = 1$ we have $P(f(B) = 1) > n^{-c_1}$.

Assume $f(B) = 1$ and therefore $\langle b_1, ..., b_m \rangle \in Y$. By the definition of Y, $\mathcal{A}$ distinguishes $\mathcal{S}$ and $\mathcal{E}$ with a bias of at least $\frac{1}{2} - n^{-1}$. Therefore the probability that $\mathcal{A}$ gives the right answer is at least $1 - n^{-1}$, in other words $P(g(B) = 1 | f(B) = 1) \geq 1 - n^{-1}$.

Assume now that $\xi = unif$. Suppose that u is fixed, $\frac{1}{2} < \|u\| \leq 1$. (Clearly the probability of this event is exponentially close to one.) We show that in this case for almost all choices of $v_1, ..., v_m$ (with an exponentially small exception) the random variables $\mathcal{S}$ and $\mathcal{E}$ are almost identical in the sense that the distance of their distribution is smaller than 2^{-n}. This will imply the required inequality, since $\mathcal{A}$ is trying to distinguish two random variables whose distance is exponentially small and which are given to it in a random order, therefore the bias of its decision is exponentially small.

We show now that if v_1,...,v_m are independent values of *unif* then with a probability exponentially close to one, we have that the distance of the distribution $\mathcal{S}_{v_1,...,v_m}$ and $\mathcal{E}_{v_1,...,v_m}$ is exponentially small.

$\mathcal{E}_{v_1,...,v_m}$ by definition has a uniform distribution on $A = \mathcal{P}^-(w_1, ..., w_n) \cap 2^{-n}\mathbb{Z}^n$. We have to show that with high probability $\mathcal{S}_{v_1,...,v_m}$ has also an almost uniform distribution on this set. (We note that this is not true if $v_1, ..., v_m$ are random values of $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$.) The elements of A form an Abelian group if we define addition as the addition in $\mathbb{R}^n$ modulo the subgroup generated by $w_1, ..., w_n$. We want to apply Lemma 3.2 for this group and the sum $\sum_{j=m/2}^{m} \delta_i v_i$. First we randomize $v_0, ..., v_{m/2}$. We show that with a probablity of at least $1 - 2^{-n}$ we have $i_0 < \frac{m}{2} - n - 1$, therefore this randomization already decides the values $w_1, ..., w_n$. To show that we estimate for a fixed j the probability of the event B_j where B_j holds iff $\mathrm{width}(v_{j+1}, ..., v_{j+n}) < a/n^2$. For any fixed $i = 1, .., n$ the probability that the distance of v_{j+i} from the subspace generated by $\{v_k | k \neq j + i\, j + 1 \leq k \leq j + n\}$ is smaller than n^{-2} is at most $\frac{2}{n\sqrt{n}}$ (see Lemma 3.4), therefore the probability that this happens for at least one i is at most $n^{-\frac{1}{2}}$. The events B_j for $j = nl$, $l = 1, ..., \frac{m}{2n} - 1$ are independent, therefore the the probability that there exists a $j = nl$, $l = 1, ..., \frac{m}{2n} - 1$ with $\neg B_j$ is at least $(1 - cn^{-\frac{1}{2}})^{\frac{m}{2n}-1} \leq 1 - 2^{-n}$ provided that $m > n^2$ and n is sufficiently large with respect to c. Therefore we may assume that $i_0 < \frac{m}{2}$ and so $w_1, ..., w_n$ are defined after the randomization of $v_1, ..., v_{m/2}$.

Definition. If $a_1, ..., a_n \in \mathbb{R}^n$ are linearly independent, and $b \in \mathbb{R}^n$, then $b \wr \mathcal{P}^-(a_1, ..., a_n)$ will denote the unique element of the half-closed prallelepiped $\mathcal{P}^-(a_1, ..., a_n)$ with the property that $b - (b \wr \mathcal{P}^-(a_1, ..., a_n))$ is an integral linear combination of the vectors $a_1, ..., a_n$.

To conclude the proof we want to show first that $\sum_{i=1}^{\frac{m}{2}} \delta_i v_i \wr \mathcal{P}^-(w_1, ..., w_n)$ is uniformly distributed on $A = \mathcal{P}^-(w_1, ..., w_n) \cap 2^{-n}\mathbb{Z}^n$, in the sense that the distance of its distribution from the uniform distribution is at most 2^{-n}. We would like to apply the following lemma, proved in [2]:

Lemma 3.2 *There exists a $c > 0$ so that if A is a finite Abelian group with n elements and k is a positive integer and $b = \langle b_1, ..., b_k \rangle$ is a sequence of length k whose elements are chosen independently and with uniform distribution from A, then with a probability of at least $1 - 2^{-ck}$ the following holds:*

Assume that b is fixed and we randomize a $0,1$-sequence $\delta_1, ...\delta_k$, where the numbers δ_i are chosen independetly and with uniform distribution from $\{0, 1\}$. For each $a \in A$ let $p_a = P(a = \sum_{i=1}^{k} \delta_i b_i)$. Then

(a) $\sum_{a \in A}(p_a - |A|^{-1})^2 \leq 2^{-2ck}$ *and*

(b) $\sum_{a \in A} |p_a - |A|^{-1}| \leq |A|^{\frac{1}{2}} 2^{-ck}$.

However, since the elements v_i, were taken not from $\mathcal{P}^-(w_1, ..., w_n)$, but from a cube, we cannot apply Lemma 3.2 directly. We will use Lemma 4.1 which is based on Lemma 4.3, a stronger version of Lemma 3.2. Applying Lemma 4.1 with $u_i \rightarrow v_{\frac{m}{2}+i-1}$, $a_i \rightarrow w_i$ (and with an appropriate change of scaling) we get that the distance of the distribution of $\sum_{i=1}^{\frac{m}{2}} \delta_i v_i \wr \mathcal{P}^-(w_1, ..., w_n)$ from the uniform distribtution is indeed at most 2^{-n}.

Since the distance from the uniform distribution is a convex function the same will be true for the sum $\sum_{i=1}^{m} \delta_i v_i$ if we are taking into account the randomization of the values $\delta_1, ..., \delta_{\frac{m}{2}-1}$.

That is, we have shown that with a probability exponentially close to one the distribution of $\mathcal{S}$ is exponentially close to the uniform distribution, that is, to the distribution of $\mathcal{E}$. *Q.E.D.*(Lemma 3.1)

Lemma 3.3 *For all c_1, c_2 there exists a c_3, c_4 and a probabilistic algorithm $\mathcal{B}$ (using an oracle) so that for all sufficiently large n, condition (1) implies condition (2), where*

(1) $\mathcal{A}$ is a probabilistic circuit with the following property: if we pick the random vector u as described in the protocol for generating the public and private keys, then with a probability of at least n^{-c_1} the following holds: $\mathcal{A}$ distinguishes the random variables unif and $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ with a bias of at least n^{-c_2},

(2) $\mathcal{B}$, using $\mathcal{A}$ as an oracle, can solve any instance of size at most n^{c_3} of the $n^{\mathcal{D}_2}$-unique shortest vector problem in time n^{c_4} and with a probability greater than $1 - 2^{-n}$.

Proof. We describe an algorithm $\mathcal{B}$ satisfying the requirements of (2). The input will be a basis of a lattice L whose shortest vector is unique upto a factor of $n^{\mathcal{D}_2}$. Let v be a shortest non-zero vector in L. Let X be the set of all $u \in \mathbb{R}^n$ so that $\frac{1}{2} \leq \|u\| \leq 1$ and $\mathcal{A}$ distinguishes the random variables *unif* and $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ with a bias of at least n^{-c_2}.

We apply Lemma 3.5 with the set X defined above. The lemma implies that if we compute $t = [n^{c_2+2}]$ (where c_2 is from Lemma 3.5) values $U_1, ..., U_t$ of the random variable ν then with a probability of at least $1 - 2^{-2n}$, at least one of the vectors $U_i v$, $i = 1, ..., t$ is in X. Assume that an i is fixed with this property. We will find the shortest vector of $U_i L$. Since $U_i = \theta V$ where $\theta \in \mathbb{R}$ and V is an orthogonal linear transformation, we have that w is an $n^{\mathcal{D}_2}$-unique shortest vector of L iff $U_i w$ is an $n^{\mathcal{D}_2}$-unique shortest vector of $U_i L$. Let J be the dual lattice of $U_i L$. Since $U_i L$ has an $n^{\mathcal{D}_2}$-unique shortest vector, J is an $(n^{\mathcal{D}_2}, 1)$ lattice in the sense of [AD]. In the remaining part of the proof we assume that the reader is familiar with this paper. To be able to apply the results of [AD] we pick a new system of coordinates so that $U_i e_j$, $j = 1, ..., n$ is the new basis. Let $K = \mathcal{K}$, that is, $KU^{(n)} = \mathcal{Q}$. As it is proved in [AD] the distance of the distributions of $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ and $\xi_{J,K,R}$ is exponentially small and clearly $\eta_K = unif$. Therefore the distinguishability of *unif*, and $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ would imply the distinguishability of $\xi_{J,K,R}$ and η_K and so the algorithm given in [AD] in polynomial time would find $J^{(d,M)}$ and so the shortest vector of $U_i L$ using $\mathcal{A}$ as an orcale with a probability exponentially close to one. *Q.E.D.*(Lemma 3.3)

Proof of Theorem 2.1. Condition (1) of the theorem is identical to condition (1) of Lemma 3.1. Its consequence condition (2) of Lemma 3.1 implies the existence of a circuit which satisfies the requirements of (1) Lemma 3.3. The conclusion of Lemma 3.3 implies the existence of an algorithm, which uses another algorithm as an oracle (the second algorithm uses a circuit as an oracle.) We can make a single algorithm from the two mentioned ones and get the conclusion of the theorem. *Q.E.D.*(Theorem 2.1)

Remark. We have shown that with high probability $i_0 < \frac{m}{2}$ only in the case when $v_0, ..., v_m$ are values of the random variable *unif*. (See the proof of Lemma 3.1.) In the same way we can also show that if $v'_1, ..., v'_m$ are random independent values of *unif* then with a probability exponentially close to one, there is an $i_1 < \frac{m}{2}$ so that $\text{width}(v'_{i_1+1}, ..., v'_{i_1+n}) > n^{-1}a$. We may get idependent values of $\mathcal{H}_{u,n^{-\mathcal{D}_1},n}$ by adding independent values of $\text{pert}_{n^{-\mathcal{D}_1},n}$ to $v'_1, ..., v'_m$. Since each possible value of $\text{pert}_{n^{-\mathcal{D}_1},n}$ is much smaller than $n^{-2}a$ we have that $\text{width}(v'_{i_1+1}, ..., v'_{i_1+n}) > n^{-1}a$ implies $\text{width}(v_{i_1+1}, ..., v_{i_1+n}) > n^{-2}a$.

Lemma 3.4 *Assume that $Q = U^n \subseteq \mathbb{R}^n$ is the unit cube of the n-dimensional space and and $H \subseteq \mathbb{R}^n$ is a hyperplane, and V is the set of those points in Q whose distance from H is at most $\gamma > 0$. Then the volume of V is at most $2\gamma n^{\frac{1}{2}}$.*

Proof. Let $b = \langle b_1, ..., b_m \rangle$ be a unit vector orthogonal to H. Since $1 = \|b\| = \sum |b_i|^2$, there is a $1 \leq j \leq n$, so that $|b_j| \geq n^{-\frac{1}{2}}$. This implies that the vectors $u, v \in V$ differ only in their jth components u_j, v_j, then $|u_j - v_j| \leq 2\gamma n^{\frac{1}{2}}$. Let Q' be the orthogonal projection of Q to the hyperspace $x_j = 0$. The previous remark implies that for for each fixed $p \in Q'$ the length of the interval in V which is projected to p is at most $2\gamma n^{\frac{1}{2}}$. Therefore the volume of Q' is at most $2\gamma n^{\frac{1}{2}}$.

Definitions. 1. We call a linear transformation U of $\mathbb{R}^n$, orthogonal, if for any $u \in \mathbb{R}^n$, $\|uU\| = \|u\|$. (An equivalent characterization of the orthogonal linear transformation U is the following: with respect to any orthonormal basis the matrix of U is orthonormal, that is, its rows as n-dimensional vectors form an orthonormal system.)

2. If the values of a random variable ξ are real numbers (or vectors, matrices with real component), then we say that a probabilistic algorithm generates ξ in polynomial time, if for any $c > 0$ there is a $c' > 0$ so that the algorithm generates a random variable in time $n^{c'}$ which approximates ξ with an error not greater then 2^{-n^c}.

Lemma 3.5 *For all $c_1 > 0$, there is a $c_2 > 0$ and a probabilistic algorithm which generates a random variable ν in polynomial time so that*

(1) each value of ν can be written in the form of $\theta\nu_1$ where $\theta \in \mathbb{R}$ and ν_1 is an orthogonal linear transformtaion of $\mathbb{R}^n$

(2) If X is a Lebesgue measurable subset of the unit ball of $\mathbb{R}^n$ whose density in it is at least n^{-c_1} and $v \in \mathbb{R}^n$ with $2^{-n^2} \le \|v\| \le 2^{n^2}$, then $P(\nu v \in X) > n^{-c_2}$.

In the proof of this lemma we will use the following well-known facts about orthogonal linear transformations. The set of all orthogonal linear transformations of $\mathbb{R}^n$ is a compact topological group under the multiplication of linear transformations and the usual topology of linear transformation (induced by e.g. any fixed matrix representation). There is a unique probability measure on this group (defined on all Borel sets) which is invariant under the mappings defined by the multiplication with any fixed element of the group. (The Haar measure of the group.) We assume that μ is a random variable taking its values with uniform distribution on the set of orthogonal linear transformations of $\mathbb{R}^n$ according to this distribtuion. We will use that following property of μ: If $v \in \mathbb{R}^n$, $\|v\| = 1$ is fixed, then μv has a uniform distribution on the set of vectors with length 1. There are several ways to generate μ in polynomial time, e.g, we may randomize sequentially the vectors $\mu e_1, ..., \mu e_n$. After $\mu e_1, ..., \mu e_i$ has been selected, μe_{i+1} is choosen with uniform distribution from the set of all unit vectors orthogonal to $\mu e_1, ..., \mu e_n$.

Let β be a random variable taking its values on the $[0, 1]$ interval and defined in the following way: first we take a vector w with uniform distribution on the unit ball of $\mathbb{R}^n$, and let $\beta = \|w\|$.

Let γ be the random variable which takes the value $(1 + \frac{1}{n})^i$ with a probability $\frac{1}{2n^4} + 1$ for $i = -n^4, ..., -1, 0, 1, ..., n^4$

Finally we assume that μ, β and γ are independent and define ν, ν_1 and θ as follows: $\nu_1 = \mu$, $\theta = \gamma\beta$, $\nu = \gamma\beta\mu$. Assume now that a $v \in \mathbb{R}^n$ is fixed with $2^{-n^2} \le \|v\| \le 2^{n^2}$. According to the definition of γ there is a γ_0 so that the probability of $\gamma = \gamma_0$ is $\frac{1}{2n^2+1}$ and $1 \le \gamma_0\|v\| \le (1 + \frac{1}{n})$.

We estimate the conditional probability $P(\nu v \in X | \gamma = \gamma_0)$. Since γ, β, μ are independent this is the (unconditional) probability $P(\gamma_0\beta\mu v \in X)$. As we have remarked earlier μv has a uniform distribution on the set of all vectors with length $\|v\|$ and so by the definition of β, $\gamma_0\beta\mu v$ has a uniform distribution on the ball around 0 with radius $\gamma_0\|v\|$. Since this ball contains the unit ball and the ratio of their volumes is at most $(1 + \frac{1}{n})^n \le 3$, we get a point in X with a probability of at least $\frac{1}{3}n^{-c_1}$, that is, $P(\nu v \in X|\gamma = \gamma_0) \leq \frac{1}{3}n^{-c_1}$ and so $P(\nu v \in X) \leq \frac{1}{3}n^{-c_1}\frac{1}{2n^2+1}$. *Q.E.D.*(Lemma 3.5).

Remarks. 1. As we have already indicated in the introduction,there are ways to make the cryptosystem more efficient. One possibility is that instead of choosing the vectors $w_1, ..., w_m$ from the set $v_1, ..., v_m$, we may randomize them separately, making sure that they are almost orthogonal, and so width$(w_1, ..., w_n)$ will be automatically large. E.g. we may pick w_i from the cube $\mathcal{K}e_i + (\mathcal{K})^{\frac{1}{2}}U^{(n)}$. The proof remains essentially the same, although it causes some additional complication that the vectors $v_1, ..., v_m, w_1, ..., w_n$ are now choosen (independently and) with $n+1$ different distributions.

4 The Distribution of $\sum_{i=1}^{\frac{m}{2}} \delta_i v_i \wr \mathcal{P}^-(w_1, ..., w_n)$

Lemma 4.1 *For all $c_1, c_2 > 0$ exists a $c_3 > 0$ so that the following holds. Suppose that n is sufficiently large, $L \subseteq \mathbb{R}^n$ is a lattice which has a basis consisting of vectors shorter than 2^{-n}, $|L \cap U^{(n)}| \leq 2^{n^{c_2}}$, where $U^{(n)}$ is the n-dimensional unit cube, $a_1, ..., a_n$ are linearly independent vectors form $U^{(n)} \cap L$ so that the width of the prallelepiped $\mathcal{P}(a_1, ..., a_n)$ is at least n^{-2}, $t \geq n^{c_3}$, and $u_1, ..., u_t$ is a random sequence of vectors taken independently and with uniform distribution from $U^{(n)} \cap L$. Then with a probability of at least $1 - 2^{-n^{c_1}}$ the sequence $u_1, ..., u_t$ has the following property:*

If we randomize a $0,1$-sequence $\delta_1, ...\delta_k$, where the numbers δ_i are chosen independently and with uniform distribution from $\{0,1\}$, then the distance of the distribution of $(\sum_{i=1}^{t} \delta_i u_i) \wr \mathcal{P}^-(a_1, ...a_n)$ from the uniform distribution on $\mathcal{P}^-(a_1, ..., a_n) \cap L$ is at most $2^{-n^{c_1}}$.

For the proof of the lemma we need a modified version of Lemma 3.2 about the distribution of subset sums in an Abelian group. The proof of Lemma 3.2 actually gives a little more then the statement of the lemma. This stronger version is formulated in Lemma 4.2 below. We do not give a proof of it, because the original proof provides this result without any modification.

Definition. Assume that ξ is a random variable, whose values are elements of the finite set A. We say that ξ is (c_1, c_2)-uniform if for each $X \subseteq A$ we have that $|X|/|A| \geq c_1$ implies $P(\xi \in X) \geq c_2$.

Lemma 4.2 *There exists a $c_1 > 0$ such that for all $c_2 > 0$ there is a $c_3 > 0$ so that the following holds. If A is a finite Abelian group with n elements and k is a positive integer and $\zeta_1, \zeta_2, ..., \zeta_k$ is a sequence of independent (c_1, c_2)-uniform random variables on the set A (with not necessarily the same distributions) and $b = \langle b_1, ..., b_k \rangle$ is a sequence of length k whose elements are random values of $\zeta_1, ..., \zeta_k$, then with a probability of at least $1 - 2^{-c_3 k}$ we have:*

Assume that b is fixed and we randomize a $0,1$-sequence $\delta_1, ...\delta_k$, where the numbers δ_i are chosen independently and with uniform distribution from $\{0,1\}$. For each $a \in A$ let $p_a = P(a = \sum_{i=1}^{k} \delta_i b_i))$. Then

(a) $\sum_{a \in A}(p_a - |A|^{-1})^2 \leq 2^{-2c_3 k}$ *and*

(b) $\sum_{a \in A} |p_a - |A|^{-1}| \leq |A|^{\frac{1}{2}} 2^{-c_3 k}$.

In the following lemma we show that if in Lemma 4.2 we replace the condition about the (c_1, c_2)-uniformity of the random variables ζ_i by a condition saying that any sum of t different random variables ζ_i is (c_1, c_2)-uniform, then the conclusion of the lemma remains true, only we have to replace k by $\frac{k}{t}$ in the exponents.

Lemma 4.3 *There exists a $c_1 > 0$ such that for all $c_2 > 0$ there is a $c_3 > 0$ so that the following holds. Assume that A is a finite Abelian group with n elements, k, t are a positive integers, $\xi_1, ..., \xi_k$ are independent random variables so that, for any sequence of integers $i_1, ..., i_t$ we have that $\eta = \sum_{j=1}^{t} \xi_{i_j}$ is (c_1, c_2)-uniform on the set A. Suppose further that $b = \langle b_1, ..., b_k \rangle$ is a sequence of length k whose elements are values of the random variables $\xi_1, ..., \xi_k$. Then with a probability of at least $1 - 2^{-c_3 \frac{k}{t}}$ we have:*

Assume that b is fixed and we randomize a $0,1$-sequence $\delta_1, ...\delta_k$, where the numbers δ_i are chosen independetly and with uniform distribution from $\{0,1\}$. For each $a \in A$ let $p_a = P(a = \sum_{i=1}^{k} \delta_i b_i))$. Then

(a) $\quad \sum_{a \in A} (p_a - |A|^{-1})^2 \leq 2^{-2c_3 \frac{k}{t}}$ *and*

(b) $\quad \sum_{a \in A} |p_a - |A|^{-1}| \leq |A|^{\frac{1}{2}} 2^{-c_3 \frac{k}{t}}$.

Proof: (b) is a consequence of (a) so we prove (a) only. Assume first that b is fixed in some arbitrary way. For the sake of simplicity we assume that k is divisible by $4t$. We partition the set $\{1, 2..., k\}$ into $k' = \frac{k}{4t}$ consecutive intervals $I_1, ..., I_{k'}$ each containing $4t$ points. We randomize the numbers δ_i, $i = 1, ..., k$ in the following way. First we partition at random each interval I_i into two subsets $J_{i,0}, J_{i,1}$. The sets $J_{i,0}$, $i = 1, .., k$ are picked independently and with uniform distribution on the set of all subsets of I_i and $J_{i,1} = I_i - J_{i,0}$. Then we randomize a $0,1$ sequence $\gamma_1, ..., \gamma_{k'}$ with uniform distribution on the set of all $0,1$ sequences of length k'. The randomization of the sequence γ_i is independent from the randomization of the sets $J_{i,0}$. Finally δ_j is defined in the following way. Assume that $j \in J_{i,\sigma_j}$, where $\sigma_j \in \{0,1\}$. Then $\delta_j \equiv \sigma_j + \gamma_j \pmod 2$, $\delta_j \in \{0,1\}$.

Let B be the event that the number of integers i with $|J_{i,0}| \geq t$ and $|J_{i,1}| \geq t$ is at least $\frac{k'}{2}$. Chernoff's inequality implies that for a fixed i $P(|J_{i,0}| \geq t \text{ and } |J_{i,1}| \geq t) \geq 1 - 2^{-c_4 t}$ where $c_4 > 0$ is an absolute constant. Therefore $P(B) \geq 1 - 2^{-c_5 k'}$ where $c_5 > 0$ is an absolute constant.

Let $\Lambda = \langle J_{1,0}, ..., J_{1,k'} \rangle$ and let Φ be the set of all sequences $Y_0, ..., Y_{k'}$, $Y_i \subseteq I_i$ with the property that there is at least $\frac{k'}{2}$ integers $i \in [1, k']$ with $t \leq |Y_i| < 3t$. (Φ is the set of all possible values of Λ with the condition B.)

The lemma states that if $\eta(a) = p_a$ and $\iota(a) = \frac{1}{|A|}$ then $\|\eta - \iota\|_{L_2} \leq 2^{-c_3 k}$, where we consider the L_2 norm according to the uniform measure μ on A, whith $\mu(A) = |A|$. (We will use later that according to this measure if $\lambda(a)$, $a \in A$ is a probability distribution on A then $\langle \lambda \rangle \leq 1$.) Let η_B resp. $\eta_{\neg B}$ be the conditional distributions of $\sum_{i=1}^{k} \delta_i b_i$ on A with the conditions B resp. $\neg B$. Clearly $\eta = P(B)\eta_B + P(\neg B)\eta_{\neg B}$. Therefore $\|\eta - \iota\| \leq \|P(B)\eta_B - \iota\| + \|P(\neg B)\eta_{\neg B}\|$. Using that $P(\neg B) \leq 2^{-c_5 k'}$ and $\|\eta_{\neg B}\| \leq 1$ we get that the second term is at most $2^{-c_5 k'}$. We may write the first term in the form: $\|(P(B) - 1)\eta_B + \eta_B - \iota\| \leq (P(B) -$

$1)\|\eta_B\| + \|\eta_B - \iota\| \leq 2^{-c_5 k'} + \|\eta_B - \iota\|$. Therefore it is sufficient to prove that $\|\eta_B - \iota\| \leq 2^{-c_6 k'}$ for some absolute constant $c_6 > 0$.

For each $X \in \Phi$ let B_X be the event $\Lambda = X$. $\eta_B = \sum_{X \in \Phi} P(B_X|B)\eta_X$, where η_X is the distribution of $\sum_{i=1}^{k} \delta_i b_i$ on A with the condition $\Lambda = X$. $\|\eta_B - \iota\| = \| \sum_{X \in \Phi}(P(B_X|B)(\eta_X - \iota)\| \leq \sum_{X \in \Phi} P(B_X|B)\|\eta_X - \iota\|$. Since $\|\eta_X - \iota\| \leq \|\eta_X\| + \|\iota\| \leq 2$ for all $X \in \Phi$ it is sufficent to show that there is an absolute constant $c_7 > 0$ so that if we take a random X on Φ with uniform distribution then with a probability of at least $1 - 2^{-c_7 k'}$ we have $\|\eta_X - \iota\| \leq 2^{-c_7 k'}$. Taking into account now the randomization of b we may say that (a) is a consequence of the following statement. There is an absolute constant c_8 so that if we randomize b and X independently and with the described distributions then with a probability of at least $1 - 2^{-c_8 k'}$ we have that $\|\eta_X - \iota\| \leq 2^{-c_8 k'}$. We prove this by showing that for every fixed $X \in \Phi$, $\|\eta_X - \iota\| \leq 2^{-c_8 k'}$ holds with a probability of at least $1 - 2^{-c_8 k'}$

Assume now that $X \in \Phi$ is fixed, $X = \langle J_{1,0}, ..., J_{k',0} \rangle$. Let W be the set of all $i = 1, ..., k'$ with the property: $|J_{i,0}| < t$ or $|J_{i,1}| < t$, let $\Theta = \bigcup_{i \in W} I_i$ and let $\Gamma = \Theta \cup \bigcup_{j=1}^{k'} J_{j,0}$. First we randomize ξ_i for each $i \in \Gamma$. Assume that we get the values b_i, $i \in \Gamma$. Now we assume that these values are fixed and we consider only the randomization of the remaining elements b_i. η_X will denote now that distribution of $\sum_{i=1}^{k} \xi_i$ with the condition $\xi_i = b_i$ for all $i \in \Gamma$. It is enough to show that for each possible fixed sequence b_i, $i \in \Gamma$,

(4.4.1) $\|\eta_X - \iota\| \leq 2^{-c_8 k'}$ holds with a probability of at least $1 - 2^{-c_8 k'}$.

Let $g = \sum_{i \in \Gamma} b_i$ and let η'_X be the distribution of $h = -g + \sum_{i=1}^{k} \delta_i \xi_i$ (with the condition $\xi = b_i, i \in \Gamma$. We will show that (4.4.1) holds with η'_X instead of η_X which is clearly equivalent to the original (4.4.1). Let W' be the complement of W in $\{1, ..., k'\}$ The definition of the numbers γ_i and δ_i imply that $h = \sum_{i \in W'} \gamma_i H_i$, where $H_i = \sum_{j \in J_{i,1}} \xi_i$. $|W|' \geq \frac{k'}{2}$, so according to Lemma 4.2 it is sufficient to show that $H_1, ..., H_{k'}$ are independent (c_1, c_2)-uniform random variables on A. This is a consequence of the following facts: (1) $\xi_1, ..., \xi_k$ are independent and the H_is are pairwise disjoint sums made from them; (2) $|J_{i,1}| \geq t$ for all, $i = 1, ..., k'$ (3) any sum formed from at least t different ξ_i is (c_1, c_2)-uniform. (We assumed that originally only for sums having exactly t terms but this more general statement easily follows by randomizing first the extra terms and fixing their values.) This completes the proof of Lemma 4.3. ■

Now we continue the proof of Lemma 4.1. According to Lemma 4.3 it is enough to show that $(\sum_{i=1}^{t} u_i) \wr \mathcal{P}$ has a (c_1, c_2)-uniform distribution on $\mathcal{P}$ for some absolute constants c_1, c_2, where $\mathcal{P} = \mathcal{P}^-(a_1, ..., a_n)$. The following lemma shows that this distribution is not only (c_1, c_2) uniform, but in a stronger sense, it is almost uniform on $\mathcal{P}$.

Lemma 4.4 *For all $c_1, c_2 > 0$ exists a $c_3 > 0$ so that the following holds. Suppose that n is sufficiently large, $L \subseteq \mathbb{R}^n$ is a lattice which has a basis consisting of vectors shorter than 2^{-n}, $|L \cap U^{(n)}| \leq 2^{n^{c_2}}$, where $U^{(n)}$ is the n-dimensional unit cube, $a_1, ..., a_n$ are linearly independent vectors form $U^{(n)} \cap L$ so that the width of the prallelepiped $\mathcal{P}(a_1, ..., a_n)$ is at least n^{-2}, $t \geq n^{c_3}$, $t \leq n^{2c_3}$ and $u_1, ..., u_t$ is a random sequence of vectors taken independently and with uniform distribution from $U^{(n)} \cap L$. Then the distance of the distribution of $(\sum_{i=1}^{t} u_i) \wr \mathcal{P}^-(a_1, ... a_n)$ from the uniform distribution on $\mathcal{P}^-(a_1, ..., a_n) \cap L$ is at most n^{-c_1}.*

Proof: Assume that c_3 is sufficiently large with respect to c_1 and c_2, $t \geq n^{c_3}$ and let $\eta' = \sum_{i=1}^{t} u_i$. If η is another distribution on L so that the distance of η and η' is at most ϵ then the distances of the distribtuions $\eta' \wr \mathcal{P}$ and $\eta \wr \mathcal{P}$ is also at most ϵ. Therefore it is sufficient to prove the lemma for any distribution η on L whose distance from η' is smaller than n^{-2c_1}.

We define such a distribution η in the following way. Let $b_1, ..., b_n$ be a basis of L so that $\|b_i\| \leq 2^{-n}$. Suppose that $g_1, ..., g_t$ is a sequence of random vectors taken independently and with uniform distribution from $U^{(n)}$, and $g = \sum_{i=1}^{t} g_i$. Let η be the unique element of L so that $g = \eta + \sum_{i=1}^{n} \gamma_i b_i$ with $0 \leq \gamma_i < 1$. In other words η is constructed by taking the sum of t independent elemetns of the cube $U^{(n)}$) (disregarding the lattice), and then finding a lattice point close to it, by using the tiling provided by the parallelepiped defined by $b_1, ..., b_n$.

First we prove that the distance of $\eta \wr \mathcal{P}^-(a_1, ..., a_n)$ from the uniform distribution on $\mathcal{P}^-(a_1, ..., a_n) \cap L$ is smaller than n^{-2c_1}, then we show that the distance of η and η' is at most n^{-2c_1}.

We will consider the distribution of η on a large cube W centered at $\langle \frac{t}{2}, ..., \frac{t}{2} \rangle$ and with sides parallel to the axis. We want to choose the length of the sides of W, so that if we tile the space with translated copies of $\mathcal{P}^-(a_1, ..., a_n)$ and $c_4 > 0$ is sufficiently large with respect to c_1 and c_2, and c_3 is sufficiently large with respect to c_4 then we have

(a) the number of parallelepipeds of the tiling completely contained in W is greater than n^{c_4} times the number or parallelepipeds only partially contained in A,

(b) if $\mathcal{P}_0$ is any fixed parallelpiped from the tiling, completely contained in W, then the distribution of η on $\mathcal{P}_0 \cap L$, has a distance at most n^{-c_4} from the uniform distribution on $\mathcal{P}_0 \cap L$,

(c) $P(\eta \in W) \geq 1 - n^{-c_4}$

These three conditions together clearly imply that $\eta \wr \mathcal{P}^-(a_1, ..., a_n)$ is sufficiently close to the uniform distribution.

Since the width of $\mathcal{P}^-(a_1, ...a_n)$ is at least n^{-2} while $\max_{i=1}^{n} \|b_i\| \leq 2^{-n}$ it is sufficient to show that (b) and (c) hold if we replace η by g in them, (we defined g in the defintion of η). (b') and (c') will denote the corresponding statements.

For the proper choice of the size of W we need the following well-known result, stating that the density function of the sum of independent random variables with uniform distribution on an interval, is very close to the density function of a normal distribution.

Lemma 4.5 *Let χ be a random variable with uniform distribution on the interval $[-1, 1]$, let χ_j be a random variable which is the sum of j independent copies of χ and let ϕ_j be the density function of χ_j. Then there exists an $\rho > 0$ so that for all sufficiently large n and for all $x \in \mathbb{R}$ we have*

$$|\frac{1}{2} n^{\frac{1}{2}} \phi_n(\frac{1}{2} n^{\frac{1}{2}} x) - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}| \leq n^{-\rho}$$

Remark. The lemma remains true if we substitute $n^{-\rho}$ by $cn^{-\frac{1}{2}}$ for a suitably chosen absolute constant c, (therefore it holds for any $\rho < \frac{1}{2}$). The reader may find a proof e.g. in V. V. Petrov, "Sums of Independent Random Variables", Springer 1975, p. 203, Theorem,

11. The more general theorem given there, is valid not only for random variables with uniform distribution, but for a wider class of random variables. We formulate the theorem below.

Let $X_1, X_2, ...$ be an infinite sequence of independent random variables with zero means and finite variances. Let $v_n(t) = E(e^{itX_n})$, $(i^2 = -1)$ and let $B_n = \sum_{j=1}^n E(X_j^2)$, $F_n(x) = P(\frac{1}{\sqrt{B_n}} \sum_{j=1}^n X_j < x)$.

Theorem 4.1 *(*Petrov, (VII.11)*). Suppose that* $\liminf \frac{B_n}{n} > 0$, $\limsup \frac{1}{n} \sum_{j=1}^n E(X_j^3) < \infty$, *and*

$$\int_{|t|>\epsilon} \prod_{j=1}^n |v_j(t)| dt = O(\frac{1}{n})$$

for every fixed $\epsilon > 0$. *Then for all sufficiently large* n *the distribution* $F_n(x)$ *has an everywhere continuous density* $p_n(x)$ *and moreover*

$$\sup_x |p_n(x) - \frac{1}{\sqrt{2\pi}} e^{-x^2/2}| = O(\frac{1}{\sqrt{n}})$$

It is easy to check that the requirements are met if each X_j is the uniform distribution on the interval $[-1, 1]$. Indeed the only non-trivial requirement is the one containing $v_j(t)$. Since X_j have a density function, say $f_j(x)$, which is $\frac{1}{2}$ on the interval $[-1, 1]$ and 0 outside, we have $v_j(t) = \int_{-\infty}^{\infty} e^{itx} f_j dx = \int_{-1}^{1} \frac{1}{2}(\cos tx + i \sin tx) dx = 2\int_0^1 \frac{1}{2} \cos txdx = \frac{\sin t}{t}$. We write the integral in the theorem as the sum of an integral for $\epsilon < |t| < 2$ and an integral for $|t| \geq 2$. The first integral is smaller than $2(\frac{\sin \epsilon}{\epsilon})^n = O(\frac{1}{n})$. The second is at most $2\int_2^{\infty} t^{-n} = O(\frac{1}{n})$.

We may conclude now the proof of Lemma 4.1. Assume that c_4 is sufficiently large with respect to c_1, c_2 and c_3 is sufficiently large with respect to c_4 and ρ, where ρ is the constant whose existence was stated in Lemma 4.5. Let $2h = 2(\log t)^{\frac{\rho}{4}} t^{\frac{1}{2}}$ be the length of a side of W. We show that the requirements (a),(b'),(c') are met.

Let $D = (\log t)^{\frac{\rho}{4}}$. Using Lemma 4.5 with $n \to t$ and $x \to D$ we get that the probability that at least one component of g is outside of the interval $[\frac{t}{2} - h, \frac{t}{2} + h]$ is at most $n(t^{-\rho} 2D + 1 - \int_{|x|<D} \frac{1}{(2\pi)^{\frac{1}{2}}} e^{-\frac{1}{2}x^2} dx) \leq 2nt^{-\frac{\rho}{4}} \leq n^{-c_4}$ which proves (c').

Since the diameter of any fixed parallelepiped $\mathcal{P}_0$ in the tiling is at most $n^{\frac{3}{2}}$ we have that the projection of $\mathcal{P}_0$ to the ith axis is an interval J_i of length at most $n^{\frac{3}{2}}$, contained in the interval $[\frac{t}{2} - h, \frac{t}{2} + h]$. Let $g = \langle \gamma_1, ..., \gamma_n \rangle$. Lemma 4.5 guarantees that there is a real number σ depending only on J_i so that for all intervals $J \subseteq J_i$ we have $\sigma |J|(1 - n^{-c_4-2}) \leq P(\gamma_i \in J) \leq \sigma |J|(1 + n^{-c_4-2})$. This implies that for any cube $Z \subseteq \mathcal{P}_0$ the probability $P(g \in Z)$ is uniquely determined by the volume of Z upto a factor of $(1 \pm n^{-c_4-2})^n$ which implies (b').

(a) is a trivial consequence of the fact that the diameter of $\mathcal{P}_0$ is at most $n^{\frac{3}{2}}$ while the diameter of W is larger than $t^{\frac{1}{2}} \geq n^{\frac{c_3}{2}}$.

Finally we have to show that the distance of η' and η is at most n^{-2c_1}. η' and η can be randomized together in the following way. We randomize η as in its original definition, using $g_1, ..., g_t$. Let b_i be a basis of L with $\|b_i\| \leq 2^{-n}$ for $i = 1, ..., n$, let u_i, $i = 1, ..., t$ be the unique element of L with $g_i = u_i + \sum_{j=1}^{n} \gamma_{i,j} b_j$, where $0 \le \gamma_{i,j} < 1$. Now we use u_i in the definition of η' instead of the original elements u_i which were chosen with uniform distribution from $U^n \cap L$. The new definition does not give exatcly the same distribution, but clearly the difference of the two distributions is exponentially small in n. So we may assume that η' is defined this way. The advantage of the new definition is that for each randomization we have that the distance of the corresponding values of η and η' in the n-dimensional space is at most $tn^2 2^{-n}$. This implies that if $H \subseteq G \subseteq L$ and G contains every point of L whose distance from H is at most $tn^2 2^{-n}$, then

(4.4.2) $P(\eta \in H) \le P(\eta' \in G)$ and $P(\eta' \in H) \le P(\eta \in G)$.

We partition W into subcubes whose sides has length between $2^{-\frac{n}{2}}$ and $2^{-\frac{n}{2}+1}$ Suppose that Θ is this partition and $T \in \Theta$. If T' resp. T'' are cubes with the same center as T, so that $T' \subseteq T \subseteq T''$, $T'^{\frac{1}{n}} = T^{\frac{1}{n}}(1 + O(n^{-c_5}))$ and $T''^{\frac{1}{n}} = T^{\frac{1}{n}}(1 + O(n^{-c_5}))$, where $c_5 < c_4/2$. Then by the already proven local uniformity of η we have that $P(\eta \in T') = P(\eta \in T)(1 + O(n^{-c_5}))$ and $P(\eta \in T'') = P(\eta \in T)(1 + O(n^{-c_5}))$. By (4.4.2) we get that $P(\eta' \in T) = P(\eta \in T)(1 + O(n^{-c_5}))$.

We complete the proof by showing that if s is a random value of η' then with an exponentially high probability we have: if $s \in T \in \Theta$ and $r = |T \cap L|$, then

(4.4.3) $P(\eta' = s) = \frac{1}{r} P(\eta' \in T)(1 + O(n^{-c_5}))$.

Since the analogue of (4.4.3) holds for η too, (by the already proven local uniformity of η) this will imply that the distance of η and η' is at most $O(n^{-c_5})$.

Proof of (4.4.3). If $u_1, ..., u_t$ are randomized as in the defintion of η' then clearly with a probability of at least $1 - 2^{-\frac{n}{3}}$, if $u_1 = \langle x_1, ..., x_n \rangle$ then for all $j = 1, ..., n$ we have $n2^{-\frac{n}{2}+1} \le |x_j| \le 1 - n2^{-\frac{n}{2}+1}$. η' induces a probability measure μ on L where $\mu(z) = P(\eta \in Z)$ for all $Z \subseteq L$. According to the previous observation there is a $B \subseteq L$ with $B \ge 1 - 2^{-\frac{n}{6}}$ so that for each fixed $s \in B$ we have $P(n2^{-\frac{n}{2}+1} \le |x_{1,j}| \le 1 - n2^{-\frac{n}{2}+1} | \eta' = s) \ge 1 - 2^{-\frac{n}{6}}$. For any fixed $s \in B$, if $d \in L$ and $u_1 + ... + u_t = s$ then $(u_1 + d - s) + u_2 + ... + u_t = d$. If $\|s - d\| \le n2^{-\frac{n}{2}+1}$ and the components of u_1 satisfy $n2^{-\frac{n}{2}+1} \le |x_j| \le 1 - n2^{-\frac{n}{2}+1}$ then $u_1 + d - s \in U^{(n)}$. We have that $P(u_1 + d - s \in U^{(n)} | \eta' = s) \ge 1 - 2^{-\frac{n}{6}}$. Therefore $P(\eta' = d) \ge P(\eta' = s) - 2^{-\frac{n}{6}}$. This implies that

(4.4.4) for any $T \in \Theta$ and $s \in T \cap B$, $d \in T \cap L$ we have $P(\eta' = d) \ge P(\eta' = s) - 2^{-\frac{n}{6}}$.

Since $\mu(B) \ge 1 - 2^{-\frac{n}{6}}$ we have that with a probability of at least $1 - 2^{-\frac{n}{12}}$, if T is the unique element of Θ containing η' then $\mu(B \cap T) \ge \mu(T \cap L)(1 - 2^{-\frac{n}{12}})$. This together with (4.4.4) implies (4.4.3). This completes the proofs of Lemma 4.4 and Lemma 4.1. ∎