Electronic Colloquium on Computational Complexity, Report No. 97 (2007)

The First and Fourth Public-Key Cryptosystems with Worst-Case/Average-Case Equivalence

Miklós Ajtai
IBM Almaden Research Center

Cynthia Dwork
Microsoft Research

October 8, 2007

Abstract

We describe a public-key cryptosystem with worst-case/average case equivalence. The cryptosystem has an *amortized* plaintext to ciphertext expansion of $O(n)$, relies on the hardness of the $\tilde{O}(n^2)$-unique shortest vector problem for lattices, and requires a public key of size at most $O(n^4)$ bits. The new cryptosystem generalizes a conceptually simple modification of the "Ajtai-Dwork" cryptosystem. We provide a unified treatment of the two cryptosystems.

1 Introduction

Since the inception of public-key cryptography it was an open problem to design a public-key cryptosystem whose security under an eavesdropping attack depends on the *worst-case*, rather than average-case, hardness of the underlying computational problem. A decade ago we provided the first such cryptosystem. The specific computational problem was the n^8-unique shortest vector problem for n-dimensional lattices, where n is a security parameter [2][1]. The scheme required transmission of vectors in $\mathbb{R}^n$, each component of which required $\Theta(n \log n)$ bits, for each bit of plaintext. This was a plaintext to ciphertext expansion factor of $O(n^2 \log n)$. Since that time many improvements in worst-case/average-case cryptosystems have been made. Two results, both due to Regev, are particularly noteworthy: (1) a public-key cryptosystem with comparable key size and blowup, but relying on the weaker complexity assumption of hardness of the $n^{1.5}$-unique shortest vector problem [14] and (2) a public-key cryptosystem having a much smaller public key ($\tilde{O}(n^2)$ instead of $\tilde{O}(n^4)$) and with a plaintext to ciphertext blowup of only $\tilde{O}(n)$, whose proof of security relies on an assumption about the *quantum* difficulty of solving the $\tilde{O}(n^{1.5})$-unique shortest vector problem for lattices [15].

In this paper we describe a cryptosystem retaining the worst-case/average-case equivalence and having linear *amortized* plaintext to ciphertext expansion: the scheme permits encoding of $\ell + 1$ plaintext bits by a single vector in $\mathbb{R}^{n+\ell}$, represented by at most $(n+\ell)cn$ bits. When $\ell \in \Theta(n)$, we achieve an amortized expansion factor of $\Theta(n)$. The semantic

[1] The n^c-*unique shortest vector* problem 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.

ISSN 1433-8092

security of the system against an eavesdropper relies on the worst-case hardness of the $\tilde{O}(n^2)$-unique shortest vector problem and requires a public key of size at most $O(n^4)$ bits. For most of this paper we assume $\ell = n$. Our proof does not rely on assumptions about quantum computing.

The new construction is an intuitive generalization of the original Ajtai-Dwork cryptosystem. Certain challenging technical obstacles to the generalization are overcome by slightly modifying the original scheme, but the fundamentals are unchanged. The current paper provides a unified presentation of both schemes.

For the reader familiar with our earlier work, we provide a brief description of the nature of the changes. Recall that in the Ajtai-Dwork cryptosystem the private key is a vector $u \in \mathbb{R}^n$ and the public key consists of a collection of points $v_1, \ldots, v_{\mathrm{poly}(n)}$ inside a large cube, close to the set of hyperplanes $\{H_i \,|\, i \in \mathbb{Z} \wedge \forall x \in H_i \langle u, x \rangle = i\}$, together with a special set of near-hyperplane points defining a parallelepiped $\mathcal{P}$. Encryptions of zero are random subset sums of the v_i, modulo $\mathcal{P}$, and encryptions of one are randomly chosen points in the interior $\mathcal{P}^-$, that is, the ciphertexts encrypting one are chosen by the sender without regard to the v_i. The intuition is that random subset sums will also be close to the hyperplane collection, while randomly chosen points in the interior $\mathcal{P}^-$ are likely to be relatively far from the hyperplane collection.

To encrypt $\ell + 1$ bits by a single vector in $\mathbb{R}^{n+\ell}$ we will use $\ell + 1$ mutually orthogonal collections of hyperplanes, now defined by $\ell + 1$ mutually orthogonal vectors $u_0, \ldots, u_\ell$. The chief algorithmic innovation in our new cryptosystem is to further refine the public key so that now, in addition to a paralellepiped $\mathcal{P}$, the key contains two sets of points: V and D. All points in the public key are now close to the intersection of all $\ell + 1$ families of hyperplanes. The $\ell + 1$ points in V, however, have a special additional structure: the $(\ell + 1) \times (\ell + 1)$ matrix A naming, for each $0 \leq i, j \leq \ell$, the specific hyperplane within each collection $\mathcal{H}_{u_j}$ to which v_i is close, will be invertible modulo 2. To encrypt an $(\ell + 1)$-bit message $b_0, \ldots, b_\ell$, the sender will compute

$$\left[\sum_i b_i v_i + \sum_j 2\delta_j d_j \right] \mod \mathcal{P}$$

where the second term is two times a random subset sum of the vectors d_j in the set D. The coefficient 2 makes the contribution of the second term "invisible" to the modulo 2 inversion process. The second term plays a crucial role in the proof of security; we argue that these random subset sums, modulo $\mathcal{P}$, are indistinguishable from uniformly chosen points in $\mathcal{P}^-$ in probabilistic polynomial time, under the assumption of worst-case hardness of the unique shortest vector problem.

The proof of semantic security in our original scheme used a hypothetical adversary's ability to dinstinguish encryptions of zeros (points close to a hyperplane collection) from ones (points chosen uniformly) to find the hyperplane $H_0 = \{x \,|\, \langle u, x \rangle = 0\}$ and, from this, the vector u. To solve the $f(n)$-unique shortest vector problem for an arbitrary lattice having an $f(n)$-unique shortest vector, the reduction first (reversibly) transforms the arbitrary lattice into a "random" lattice with an $f(n)$-unique shortest vector u, and then examines the dual of this transformed lattice. The points in the dual lattice are arranged on the hyperplanes induced by u, just as the points in the public key of the cryptosystem are on the hyperplanes induced by the private key. By perturbing randomly chosen points in the dual lattice one obtains something looking very much like a public key of the cryptosystem. We then used the aforementioned capability of the adversary to discover H_0, and hence the unique-shortest vector u in the transformed lattice. A shortest vector in the original lattice is obtained by reversing the transformation.

The proof of the new scheme uses the same idea, but must now take points close to zero, respectively one, collections of $(n-1)$-dimensional hyperplanes in $\mathbb{R}^n$, and "lift" them to higher dimensions, where they will be close to some number $k \leq \ell$, respectively, $k+1 \leq \ell+1$ mutually orthogonal collections of hyperplanes, in order to get something that looks like a public key in the new system. This introduces technical difficulties in the sampling, as well as in the perturbations, which originally are in R^n and must now be in a higher dimensional space. These difficulties are addressed by sampling and perturbing according to high-dimensional spherical Gaussian distributions. Because the coordinates of the Gaussian operate independently, lifting a point in $\mathbb{R}^n$ close to a single family of hyperplanes to, say, a point in $\mathbb{R}^{n+1}$ close to two mutually orthogonal families of hyperplanes, is intuitively straightforward.

Finally, we rely on results of Regev [14] and of Regev and Micciancio [13], as well as a variety of minor technical tricks, to obtain the $\tilde{O}(n^2)$ approximation factor for the unique shortest vector problem.

2 Preliminaries

Given m linearly independent vectors $(b_1, \ldots, b_m) \in \mathbb{R}^m$, the lattice $L = L(b_1, \ldots, b_m)$ is the set of all integer linear combinations of the basis vectors:

$$L = \{\alpha_1 b_1 + \ldots + \alpha_m b_m \mid \alpha_1, \ldots, \alpha_m \in \mathbb{Z}\}.$$

The *dual* of L, frequently denoted L^*, is the lattice with basis vectors $c_1, \ldots, c_m$, where for all $1 \leq i, j \leq m$: $\langle c_i, b_j \rangle = \delta_{ij}$. Here $\delta_{ij} = 0$ if $i \neq j$ and 1 otherwise.

Rounding. We round everything to $\mathbf{p}$ bits of precision. That is, for $x \in \mathbb{R}$, define $\text{Round}_\alpha(x) = i\alpha$, where i is the largest integer with $i\alpha \leq x$. If $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ then $\text{Round}_\alpha(x) = (\text{Round}_\alpha(x_1), \ldots, \text{Round}_\alpha(x_m))$. We choose $\alpha = 2^{-\mathbf{p}}$. Thus, when we sample from real-valued distributions we always mean that we are taking rounded values.

Norms. For $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$, $\|x\|$ denotes the L_2 norm of x: $\|x\| = (\sum_{i=1}^m x_i^2)^{1/2}$. We use $\|x\|_\infty$ to denote the L-infinity norm: $\|x\|_\infty = \max_{i=1}^m |x_i|$.

Negligible. A function $\nu(n)$ that is asymptotically smaller than the inverse of any polynomial in n is said to be *negligible in n*.

Notation ($\mathcal{N}_m(0, \sigma^2)$, $g_\sigma^m(y)$). For all integers $m \geq 1$ and reals $\sigma > 0$ we let $\mathcal{N}_m(0, \sigma^2)$ denote the m-dimensional spherical Gaussian distribution centered at the origin, given by the probability density function

$$g(y) = \frac{1}{(\sigma\sqrt{2\pi})^m} e^{-\frac{||y||_2^2}{2\sigma^2}}$$

for $y \in \mathbb{R}^m$. We may refer to $\sqrt{m}\sigma$ as the *radius* of $\mathcal{N}_m(0, \sigma^2)$ and σ as the *parameter* of the Gaussian. So when $m = 1$ the radius is just the standard deviation, and in general the radius is the expected length of a sample from the distribution. When we wish to emphasize the dimension m and the parameter σ of the probability density function we may write: $g_\sigma^m(y)$. Letting $a_1, \ldots, a_m$ be any orthornomal basis for $\mathbb{R}^m$, $\mathcal{N}_m(0, \sigma^2)$ may be sampled by independently choosing values $n_1, \ldots n_m$ of the 1-dimensional normal with mean 0 and variance σ^2, (ie, m independent values of $\mathcal{N}_1(0, \sigma^2)$); the result is the sum $\sum_{i=1}^m n_i a_i$.

Notation ($\mathbf{e_{u_i}}, z_i, \mathcal{H}_u, \mathcal{H}$). The private key will be a collection of $\ell+1$ mutually orthogonal vectors $u_0 \ldots, u_\ell$, each of length close to unity. For each u_i, $i = 0, \ldots, \ell$, we define two vectors: $\mathbf{e_{u_i}} = u_i/\|u_i\|$ is the unit vector parallel to u_i and $z_i = u_i/\|u_i\|^2 = \mathbf{e_{u_i}}/\|u_i\|$ is the vector parallel to u_i satisfying $z_i \cdot u_i = 1$, so $||z_i|| = 1/||u_i||$. Sometimes we have a single u, in which case we let $\mathbf{e_u} = u/\|u\|$. For any m, a vector $u \in \mathbb{R}^m$ implicitly yields a collection $\mathcal{H}_u$ of $(m-1)$-dimensional hyperplanes as follows: for any integer i, the ith hyperplane is the set of points in $\mathbb{R}^m$ whose inner product with u equals i. Abusing notation and simply writing $\mathcal{H}_u$ for $\cup\mathcal{H}_u$, we let $\mathcal{H}'$ denote the intersection of the hyperplanes induced by the $\ell + 1$ vectors $u_0, \ldots, u_\ell$ of the private key: $\mathcal{H} = \cap_{i=0}^{\ell}\mathcal{H}_{u_i}$.

Definition (the distribution $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_\ell}^{m,R}$). The superscript m denotes the dimension of the space, so $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_\ell}^{m,R}$ is a distribution on vectors in $\mathbb{R}^m$; R is a scalar, and the u_i are vectors in $\mathbb{R}^m$. We first define the distribution $\mathcal{H}\mathrm{Samp}_u^{1,R}$ by the following experiment. Choose i' by sampling from $\mathcal{N}_1(0, R^2)$. Output $i = \mathrm{Round}_{1/\|u\|}(i')$. Intuitively, $\mathcal{H}\mathrm{Samp}_u^{1,R}$ will be used to chose a hyperplane, from $\mathcal{H}_u$, with a "Gaussian-like" distribution, provided R is sufficiently large. (This is made formal in Lemma 5.1 and Claim 5.1 below.) For any integer $\ell \geq 0$, the distribution $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_\ell}^{m,R}$ is defined by the following process. For $j = 0, \ldots, \ell$, let $i_0, \ldots, i_\ell$ be independent samples of $\mathcal{H}\mathrm{Samp}_{u_j}^{1,R}$, respectively. Let $y = (y_1, \ldots, y_{m-(\ell+1)})$ be drawn from $\mathcal{N}_{m-(\ell+1)}(0, R^2)$. Let $f_1, \ldots, f_{m-(\ell+1)}$ be an orthonormal basis for $\mathrm{span}(u_0, \ldots, u_\ell)^\perp$. Output

$$\sum_{j=0}^{\ell} i_j \mathbf{e_{u_j}} + \sum_{k=1}^{m-(\ell+1)} f_k y_k$$

Note that for each $j = 0, \ldots, \ell$ we have that i_j is an integer multiple of $1/\|u_j\|$, and so $i_j\mathbf{e_{u_j}}$, which is the projection of the output onto $\mathrm{span}(u_j)$, has length which is a multiple of the distance between adjacent hyperplanes in $\mathcal{H}_{u_j}$.

Notation ($\xi_\mathcal{K}, \pi_\rho$). We will define two parameterized random variables: ξ, parameterized by $\mathcal{K}$, and π, parameterized by ρ.

Definition of the random variable ξ. The random variable $\xi_{\mathcal{K}}$ will take on values in the intersection $\mathcal{H} = \cap_{i=0}^{\ell} \mathcal{H}_{u_i}$. Specifically, we will take $\xi_{\mathcal{K}} = \mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_\ell}^{m,\mathcal{K}(n)}$, where $m = n + \ell$.

Definition of the random variable π. The random variable π will be a called a *perturbation*; it has an m-dimensional spherical Gaussian distribution with a small radius. Letting $m = n + \ell$, we define π_ρ to be $\mathcal{N}_m(0, \rho^2)$; the radius, or expected length of a sample from $\mathcal{N}_m(0, \rho^2)$ is $\sqrt{m}\rho$.

The public key will consist of several vectors in $\mathbb{R}^{n+\ell}$, each an independent value of the random variable $\xi_{\mathcal{K}} + \pi_\rho$, for independent ξ and π. For now we note that if $v_0, \ldots, v_\ell$ are values of the random variable $\xi_{\mathcal{K}} + \pi_\rho$ then the matrix whose rows are the vectors $v_1, \ldots, v_\ell$ may be written as $(AZ + Q) + T$ where

- Q is a matrix whose rows are random values of π_ρ;
- A is integer-valued;
- Z is the $(\ell + 1) \times (n + \ell)$ matrix with rows $z_0, \ldots, z_\ell$ (recall that for $i = 0, \ldots, \ell$ the vector z_i is parallel to u_i and satisfies $z_i \cdot u_i = 1$); and
- T is a matrix whose rows are vectors in the $(n - 1)$-dimensional space orthogonal to $\mathrm{span}(z_0, \ldots, z_\ell)$.

Note that AZ and Q are independent and that T depends only on $z_0, \ldots, z_\ell$.

3 The Compact Cryptosystem

We describe the components of the cryptosystem: private key generation, public key generation, encryption, and decryption.

The Private Key. The private key is a collection of $\ell + 1$ mutually orthogonal vectors $u_0, \ldots, u_\ell$, each of length close to unity. We construct the key as follows. For each $i = 0, 1, \ldots, \ell$, we choose a length for u_i by independently sampling $\mathcal{B}^{(n)}(0, 1)$, the n-dimensional ball centered at the origin with radius 1, and taking the length of the sample. Given that we have chosen $u_0, \ldots, u_{i-1}$, we choose a direction for u_i in $\mathrm{span}(u_0, \ldots, u_{i-1})^\perp$ by independently sampling from $\mathcal{N}_{n+\ell-i}(0, 1)$. Equivalently, the private key is the $(\ell+1)\times(n+\ell)$ matrix Z defined above, with rows $z_i = u_i/\|u_i\|^2$, for $i = 0, \ldots, \ell$.

The Public Key $(\mathcal{P}, V, D)$. The public key consists of three components: a parallelepiped $\mathcal{P}$, defined by $n+\ell$ vectors $p_1, \ldots, p_{n+\ell}$, and two ordered sets of vectors $V = \{v_0, \ldots, v_\ell\}$ and $D = \{d_1, \ldots, d_{m'}\}$. Roughly speaking, $\mathcal{P}$ plays a technical role in the encryption process, serving as a "bounding region" for ciphertexts. The set V will be used for encoding $(\ell + 1)$-bit messages; however, this encoding will not be semantically secure. The set D will, intuitively, be used for creating "dust" to convert the encoding into a secure encryption. We now describe how each of the components is chosen.

The Choice of $\mathcal{P}$. We repeatedly choose independent values $x_1, \ldots, x_{n+\ell}$ of the random variable $\xi_{\mathcal{K}} + \pi_\rho$, where ξ and π are independent as well, until $x_1, \ldots, x_{n+\ell}$ define a parallelepiped of width[2] at least $\mathcal{K}(n)/(n+\ell)^2$ (this occurs with at least constant probability; see Lemma 8.2). We then set $p_i = x_i$, $i = 1, \ldots, n+\ell$. Finally, we perturb the parallelepiped very slightly so that the number of points in $\mathcal{P}^- \cap 2^{-\mathbf{P}}\mathbb{Z}^{n+\ell}$ is odd. This is done as follows: if $\mathcal{P}$ (after rounding) has even volume, then let P be the matrix with columns $p_1, \ldots, p_{n+\ell}$. We modifiy the least significant bits of each entry in P so that the diagonal entries are odd and the off-diagonal entries are even. The modified matrix has odd determinant. Its columns are the vectors of the modified $\mathcal{P}$.

The Choice of V. Assume the set $\mathcal{P}$ has already been chosen. The $\ell + 1$ vectors in $V = (v_0, \ldots, v_\ell)$ will be independently chosen values of the random variable $\xi_{\mathcal{K}} + \pi_\rho$, where ξ and π are independent as well. We require that the set $V = (v_0, \ldots, v_\ell)$ satisfy the following constraint: when we express the $(\ell+1) \times (n+\ell)$ matrix with rows $v_0, \ldots, v_\ell$ as the sum of $(AZ+Q)$ and T, as discussed above, then A must be invertible modulo 2. This will happen with constant probability (see Lemma 8.1). We therefore repeatedly choose a fresh set $V = (v_0, \ldots, v_\ell)$ until this invertibility condition is satisfied.

The Choice of D. The set $D = (d_1, \ldots, d_{m'})$ consists of m' independently chosen values of the random variable $\xi_{\mathcal{K}} + \pi_\rho$, where ξ and π are also independent. As we will see, m' will be roughly of size $O(n^2)$. This is the dominant factor in determining the size of the public key.

Remark 3.1 *If one of the perturbations (values of the random variable π_ρ) is very large (an event of negligible likelihood), the public key may be problematic. Because this is so unlikely we simply ignore the possibility. When needed, we state explicit assumed upper bounds on the sizes of the perturbations.*

Encryption. The message $b_0 b_1 \ldots b_\ell$ is encrypted as follows: Choose $\delta_1, \ldots, \delta_{m'}$ independently so that each $\delta_i \in_R \{0,1\}$, the ciphertext is:

$$\left[\sum_{i=0}^{\ell} b_i v_i + \sum_{i=1}^{m'} 2\delta_i d_i\right] \bmod \mathcal{P} \ .$$

There will be no decryption error.

Before describing the decryption process we give some intuition for the encryptions. Recall that the $(\ell+1) \times (n+\ell)$ matrix with rows $v_0, \ldots, v_{m'}$ can be expressed as the sum of $(AZ+Q)$ and T, where Q is a matrix whose rows are random perturbations, the $(\ell+1) \times (\ell+1)$ matrix A is invertible modulo 2, Z is the $(\ell+1) \times m$ matrix with rows $z_0, \ldots, z_\ell$, and where T is a fixed matrix whose rows span the $(m-\ell-1)$-dimensional space orthogonal to span$(z_1, \ldots, z_{\ell+1})$.

Given Z, A, T and an encoding $\sum_{i=0}^{\ell} b_i v_i \bmod \mathcal{P}$ of $b_1 b_1 \ldots b_\ell \in \{0,1\}^{\ell+1}$, it is possible to project out the components in T and then solve for $b_0, \ldots, b_\ell$. However, such an encoding

[2] The width of a parallelepiped defined by linearly independent vectors $x_1, \ldots x_m$ is the maximum of the distances between the point x_i and the subspace generated by the remaining $\{x_j \mid j \neq i\}$, for $i = 1, \ldots, m$.

would not be semantically secure; for example, it would be easy to determine if the encrypted message is equal to any given value of $b'_0 b'_1 \dots b'_\ell$. To make the encoding semantically secure we add to it a random subset sum of the vectors $d_1, \dots, d_{m'}$ – multiplied by 2. We will prove that such a random subset sum is indistinguishable from a randomly chosen point in $\mathcal{P}^-$. (Since $2^{-\mathbf{p}}\mathbb{Z}^{n+\ell} \cap \mathcal{P}^-$ contains an odd number of points, $2x \bmod \mathcal{P}^-$ has the same distribution as x.) Thus, still intuitively, the final ciphertext is indistinguishable from a point chosen uniformly from $2^{-\mathbf{p}}\mathbb{Z}^{n+\ell} \cap \mathcal{P}^-$. During decryption the randomizing value disappears if we mod out by 2 after removing the perturbation, which requires knowing Z, the private key. Since A is invertible modulo 2 this operation leaves the sequence $b_0 \dots b_\ell$ intact.

Ciphertext Decryption. Let $m = n + \ell$. For $i = 1, \dots, m'$, let $t(d_i)$ be the projection of d_i onto the row space of T, and for $i = 0 \dots, \ell$ let $t(v_i)$ be the projection of v_i onto the row space of T. Then, letting x be the ciphertext to be decrypted, we have

$$x = \left[\sum_{i=0}^{\ell} b_i v_i + \sum_{i=1}^{m'} 2\delta_i d_i \right] \bmod \mathcal{P},$$

where the plaintext message to be extracted is $b_0 \dots b_\ell$. From the definition of the d_i and v_i we have

$$\begin{aligned} x &= \left[\sum_{i=0}^{\ell} b_i [(\sum_{j=0}^{\ell} \alpha_{ij} z_j) + t(v_i) + \mathcal{N}_m(0, \rho^2)] \right. \\ &+ \left. \sum_{i=1}^{m'} 2\delta_i [(\sum_{j=0}^{\ell} \beta_{ij} z_j) + t(d_i) + \mathcal{N}_m(0, \rho^2)] \right] \bmod \mathcal{P} \end{aligned}$$

for some appropriate α_{ij}, where we have ensured that

$$A = (\alpha_{ij})_{i,j=0,\dots,\ell}$$

is invertible modulo 2, and some β_{ij} (not of interest to us here) . The decryption process comprises four steps:

1. Change basis so as to express x as a linear combination of $z_0, \dots, z_\ell$ and a basis for T, and project out the components in T; call the result $x^{(1)}$.

2. Round the entries of $x^{(1)}$ to remove the perturbations; call the result $x^{(2)}$. Specifically, for $1 \le i \le \ell$, the inner product $u_i \cdot x^{(1)}$ is rounded to the *nearest* integer. Thus

$$x_i^{(2)} = \begin{cases} \text{round}(u_i \cdot x^{(1)}) & \text{if } 0 \le i \le \ell \\ 0 & \text{if } \ell + 1 \le i \le n + \ell \end{cases} \tag{1}$$

   Note that $x^{(2)} \in \mathbb{Z}^{n+\ell}$.

3. Compute $x^{(3)} = x^{(2)} \bmod 2$, thereby eliminating the multiples of $d_1, \dots, d_{m'}$.

4. Solve the resulting system of linear equations (modulo 2) to obtain $b_0, \ldots, b_\ell$. This is where we use the fact that A is invertible modulo 2.

This completes the description of the cryptosystem.

Lemma 3.1 *Let $m' = n^2 \log^5 n$. There exists a constant $c > 0$ such that for all functions $\mathcal{K}(n) \geq 2^{cn}$ and $\rho = \rho(n) \leq 1/(2n^{4.5} \log^8 n)$, if the public key contains no perturbation of infinity norm greater than $\log^2 n\rho$ (that is, during the process of generating the perturbation no component had length exceeding this bound) and no vectors of magnitude greater than $\log^2 n(\sqrt{n+\ell})\mathcal{K}(n)$, then every encryption is decrypted correctly.*

Note that the conditions on the perturbations π_ρ and vectors $\xi_{\mathcal{K}(n)}$ will hold with probability all but negligible in n.

Proof: Suppose for a moment that $\rho = 0$, so that there is no perturbation, and all points in the public key are in $\mathcal{H}$. In this case correct decryption follows immediately from the description of the encryption and decryption processes. However, when $\rho > 0$, as required for the proof of security below, difficulties arise from two sources.

Since the hyperplane collections are mutually orthgonal, and since the perturbations π_ρ are drawn from spherically symmetric Gaussians, we may confine our attention to a single family $\mathcal{H}_u$ of hyperplanes; the following analysis applies to all ℓ+1 collections simultaneously. By assumption, $\mathrm{dist}(v_i, \mathcal{H}_u) \leq \log^2 n\rho$ for $1 \leq i \leq \ell + n$. A similar statement applies to the d_j, for $1 \leq j \leq m'$. For the correct outcome in the rounding step we need that $\mathrm{dist}(\sum_i b_i v_i + \sum_j 2\delta_j d_j, \mathcal{H}_u) \leq 1/2\|u\|$. Under the assumptions of the lemma we get an upper bound on this distance of $(2m' + \ell + 1)\log^2 n\rho$. Since $\|u\| \in [1/2, 1]$, with no mitigating arguments this imposes a restriction of $\rho < 1/2(2m' + \ell + 1)\log^2 n$. As we will see (Lemma 7.2), we may take $m' = n^2 \log^5 n$.

In addition, the vertices of the parallelepiped $\mathcal{P}$ are themselves not in the hyperplane collections, only close to them. If the ciphertext, before modding out by $\mathcal{P}$, is of the form $\sum_{i=1}^{n+\ell} \lambda_i p_i + \mathbf{o}$, where $\mathbf{o} \in \mathcal{P}^-$ and $\lambda_1, \ldots, \lambda_{n+\ell} \in \mathbb{Z}$, then the distance from the hyperplane collection contributed by the modding out step in the worst case can be as large as

$$\sum_{i=1}^{n+\ell} \lambda_i \mathrm{dist}(p_i, \mathcal{H}_u).$$

We therefore need to bound the λ_i. Let $w = \mathcal{K}(n)/(n+\ell)$, the lower bound on the width of $\mathcal{P}$. For any $C \in \mathbb{R}$, consider a point y at distance C from the origin. We claim that if we write y as $\sum_{i=1}^{n+\ell} \lambda_i p_i + \mathbf{o}$, for integer λ_i, then $\forall i\, \lambda_i \leq \lfloor C/w \rfloor + 1$. To see this, consider the ray from the origin to y, and divide it into pieces of length w (except possibly the last piece, which may be shorter). Tile the space $\mathbb{R}^{n+\ell}$ with copies of $\mathcal{P}$. As we move from the origin towards y along one of the pieces, we may move between adjoining tiles, but since each piece is of length w we can increase each λ_i by at most 1 in this process. Since there are at most $\lfloor C/w \rfloor + 1$ pieces in the segment the claim follows.

Now, by assumption each $x \in V \cup D$ is of length at most $\log^2 n\sqrt{n+\ell}\,\mathcal{K}(n)$. Thus the ciphertext, before modding out by $\mathcal{P}$, has length at most

$$\begin{aligned}(2m' + \ell + 1)\log^2 n\sqrt{n+\ell}\,\mathcal{K}(n) &= (2m' + \ell + 1)\log^2 n\sqrt{n+\ell}\,w(n+\ell) \\ &= (2m' + \ell + 1)\log^2 n(n+\ell)^{3/2} w.\end{aligned}$$

Taking $\ell = n$ and $m' = n^2 \log^5 n$, when n is sufficiently large this is bounded by $n^{3.5}(\log^8 n)\, w$. Thus, before modding out by $\mathcal{P}$, if we write the ciphertext as $\sum_{i=1}^{n+\ell} \lambda_i p_i + \mathbf{o}$, we get an upper bound of $n^{3.5} \log^8 n$ for each λ_i, or $n^{4.5} \log^8 n$ for $\sum_i \lambda_i$.

Setting $\rho < 1/(2n^{4.5} \log^8 n \|u\|)$ covers both sources of distance from each hyperplane collection and ensures correct rounding. ■

Following standard terminology, the security parameter for the cryptosystem is n.

We now describe the type of security offered by the cryptosystem against an eavesdropping adversary: indistinguishability of encryptions under the worst-case hardness assumption of the $f(n)$-unique shortest vector problem for lattices.

Definition ($f(n)$-unique shortest vector problem for lattices). Let $L \subseteq \mathbb{Z}^n$ be a lattice. Assume that L contains a nonzero vector u such that the shortest non-zero vector $v \in L$ not parallel to u has length at least $f(n)||u||$. Then we say that u is an $f(n)$-unique shortest vector for L. The $f(n)$-unique shortest vector problem for lattices is to find an $f(n)$-unique shortest vector for $L \subseteq \mathbb{Z}^n$, when L has such a vector.

The literature contains several equivalent definitions of semantic security against an eavesdropping adversary. We have chosen to work with one of those based on indistinguishability of encryptions:

Definition (indistinguishabilty of encryptions against an eavesdropping adversary). We use a standard definition of indistinguishability of a cryptosystem against an eavesdropper [9]. We must describe the nature of the eavesdropping attack and what it means to break the system.

The Eavesdropping Attack. The adversary is given a public key generated with security parameter n; the adversary produces a pair of $(\ell + 1)$-bit messages. The adversary is then given an encryption, under the same key, of one of the two messages, where each message is selected for encryption with probability 1/2. The adversary outputs a guess as to which of the two messages was encrypted.

Indistinguishability. The system is secure if for all probabilistic polynomial time adversaries the probability, over all the random choices – the private and public keys, randomness used by the adversary, randomness used in selecting which message to encrypt, and the randomness used during the encryption process – that the adversary is correct is negligibly close to 1/2.

Thus, if the system is insecure, then there exists a probabilistic polynomial time bounded adversary $\mathcal{C}$ and a polynomial $P_1(n)$, such that on at least a $1/P_1(n)$ fraction of the keys $\mathcal{C}$ chooses with probability at least $1/P_1(n)$ a pair of messages on which it guesses correctly with probability at least $1/2 + 1/P_1(n)$. For ease of notation, we let $p_1(x)$ denote $1/P_1(x)$ for $x \in \mathbb{R}$.

Theorem 3.1 *There exists an $f(n) \in \tilde{O}(n^{4.5})$ such that if the compact cryptosystem is insecure against an eavesdropping adversary then there is a probabilistic polynomial time bounded algorithm that takes as input an arbitrary lattice $L \subseteq \mathbb{Z}^n$ with an $f(n)$-unique* *shortest vector, presented by a basis where the logarithm of the length of each basis vector is polynomial in n, and outputs an $f(n)$-unique shortest vector for L with probability polynomial in n. The probability is over the coin flips of the algorithm.*

4 Proof of Indistinguishability of Encryptions: Overview

We show that if the system is not semantically secure with respect to chosen plaintext attacks, then we can solve arbitrary instances of the $f(n)$-unique shortest vector problem. The structure of the argument is as follows:

1. Assume the scheme is insecure, and let $\mathcal{C}$ be a probabilistic polynomial time bounded adversary that breaks the system.

2. Observe that if the points in the public key are chosen exactly as described above but without regard to whether or not A is invertible modulo 2, then there is a constant probability p_A that the resulting public key is valid. This is proved in Lemma 8.1.

3. Show (Lemma 7.3) that if the points in the "public key," are chosen by first sampling from $\mathcal{N}_{n+\ell}(0, \mathcal{K}(n)^2)$ (but not close to any hyperplane collection) and then perturbing according to $\mathcal{N}_{n+\ell}(0, \rho^2)$, then the adversary $\mathcal{C}$ will guess correctly with probability negligibly (in n) close to 1/2. Specifically, we prove that if D is chosen according to the distribution
$$D \in_R [\mathcal{N}_{n+\ell}(0, \mathcal{K}(n))^2) + \mathcal{N}_{n+\ell}(0, \rho^2)]^{m'}$$
then with all but negligible probability over the choice of D, random subset sums $\sum_{i=1}^{m'} \delta_i d_i \bmod \mathcal{P}$ have statistical distribution negligibly close to the uniform distribution modulo $\mathcal{P}$.

4. Design a polynomial time bounded distinguisher $\mathcal{D}$ that, using $\mathcal{C}$ as an oracle, distinguishes between "public keys" containing points drawn from $\mathcal{H}\text{Samp}_{u_0,\ldots,u_\ell}^{n+\ell,\mathcal{K}(n)} + \mathcal{N}_{n+\ell}(0, \rho^2)$ and points chosen from $\mathcal{N}_{n+\ell}(0, \mathcal{K}(n)^2) + \mathcal{N}_{n+\ell}(0, \rho^2)$ with polynomial advantage. One such distinguisher is described below. The distinguisher will accept with probability at least $\frac{1}{2} + p_A p_1^2(n)/2$ in the first case and at most $\frac{1}{2} + \nu(n)$ in the second case. The probability space for the distinguisher is over everything: the choice of private key, public key, coin flips of the distinguisher, and the random choices of $\mathcal{C}$. The factor p_A arises because the assumption in Step 1 only applies when the public key is valid.

5. Conclude via a hybrid argument that for some $k \in \{0, \ldots, \ell - 1\}$ there is a drop from, say, q_H to $q_L \leq q_H - p_A p_1^2(n)/2(\ell + 1)$, in the probability that the distinguisher $\mathcal{D}$ accepts; that is, it accepts with probability q_H when given "public keys" of points chosen close to $k + 1$ families of hyperplanes, but only q_L on "public keys" close to only k families of hyperplanes. All the points are in the full space $\mathbb{R}^{n+\ell}$.

   In a little more detail: at every step in the chain a private key $u_0, \ldots, u_\ell$ is chosen. At one extreme (Step 0 in the chain) the points in the public key are chosen without regard to the families of hyperplanes induced by the vectors in the private key. At the kth step in the chain the points in the public key are chosen to be close to the families induced by $u_0, \ldots, u_{k-1}$, formally, according to $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_{k-1}}^{n+\ell,\mathcal{K}(n)} + \mathcal{N}_{n+\ell}(0, \rho^2)$. In the final element in the chain the points $(\mathcal{P}, V, D)$ of the public key are selected to be close to all $\ell+1$ families of hyperplanes, formally, according to $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_\ell}^{n+\ell,\mathcal{K}(n)} + \mathcal{N}_{n+\ell}(0, \rho^2)$. These points define a valid public key when the matrix A is invertible modulo 2, which happens with constant probability p_A. The probability space for the distinguisher is over the choice of $u_0, \ldots, u_{k-1}$, the remaining randomness in generating the public key, choices made during the encryption, and the coin flips of $\mathcal{C}$. Note that when $k = 0$ there is no dependence on the private key at all.

Description of the Distinguisher $\mathcal{D}$. Given a putative public key E, the distinguisher has $\mathcal{C}$ generate a pair of messages m_0, m_1. It then tests the adversary on encryptions of these messages under E to determine if this is a "good" case for $\mathcal{C}$. To do this, it runs the following experiment a number $N \gg P_1^2(n)$ of times: Randomly choose $i \in_R \{0, 1\}$, and create a ciphertext $\alpha \in_R E(m_i)$. Give α to $\mathcal{C}$, which outputs a guess of i. The experiment succeeds if $\mathcal{C}$ guesses correctly on at least $N(1/2+p_1(n)/2)$ tries. If the experiment succeeds, then $\mathcal{D}$ accepts; otherwise, it flips a fair coin and accepts if the coin comes up heads, rejecting otherwise.

We assume without loss of generality that the parallelepiped in $E = (\mathcal{P}, V, D)$ has sufficient width, since no secret information is needed to check this. At the top level of the chain, E will be a valid public key with probability at least p_A. Conditioned on the public key being valid, the assumption of insecurity implies that with probability at least $p_1(n)^2$ this is a "good" case for $\mathcal{C}$, in the sense that $\mathcal{C}$ can distinguish encryptions of m_0 and m_1 with advantage at least $p_1(n)$. Let X_i, $i = 1, \ldots, N$, be a random variable with value 1 if $\mathcal{C}$ guesses correctly on the ith trial, and 0 otherwise. Then, in this good case, $\Pr[X_i = 1] \geq 1/2+p_1(n)$, and, letting $X = \sum_i X_i$ we have $\mu = E[X] = N(1/2 + p_1(n))$. Using the Cherneoff bound $\Pr[X < (1 - \delta)\mu] < \exp(-\mu\delta^2/2)$ we get that in this good case $\Pr[X < N/2 + Np_1(n)/2] < \exp(-N(p(n))^2/4)$. This follows from the following simple calculation. Let $z = N(1/2 + p_1(n)/2)$; write $z = (1 - \delta)\mu$ and solve for δ:

$$\begin{array}{rcl}
N(1/2 + p_1(n)/2) & = & (1 - \delta)N(1/2 + p_1(n)) \\
1/2 + p_1(n)/2 & = & (1 - \delta)(1/2 + p_1(n)) \\
1/2 + p_1(n)/2 & = & 1/2 + p_1(n) - \delta(1/2 + p_1(n)) \\
\delta(1/2 + p_1(n)) & = & p_1(n)/2 \\
\delta & = & p_1(n)/(1 + 2p_1(n)) > p_1(n)
\end{array}$$

So $\Pr[X < (1-\delta)\mu] < \Pr[X < (1-p_1(n))\mu]$ and by the Chernoff bound this is at most $\exp(-\mu(p_1(n))^2/2)$. Since $\mu > N/2$ this is at most $\exp(-N(p_1(n))^2/4)$. Thus, taking N sufficiently large, say, $N = (P_1(n))^3$ ensures an exponentially small probability of a failed experiment in the good case.

If we are not in a "good" case for $\mathcal{C}$ then, regardless of whether or not A is invertible modulo 2, we only know that the probability of success is at least 1/2. It follows that at the top level of the chain the probability that the distinguisher $\mathcal{D}$ accepts is at least

$$p_A(p_1(n))^2(1-\nu(n)) + (1-p_A(p_1(n))^2)\frac{1}{2} = \frac{1}{2} + p_A(p_1(n))^2/2 - p_A(p_1(n))^2\nu(n)$$

On the other hand, when the points in E are drawn from $\mathcal{N}_{n+\ell}(0, \mathcal{K}(n)^2)$, that is, at Step 0 of the chain, the probability that the adversary guesses correctly is negligibly close to 1/2. Letting X_i, $i = 1, \ldots, N$, be a random variable that takes value 1 when the adversary guesses correctly and 0 otherwise, and letting $X = \sum_i X_i$, the probability of a successful trial is the probability that X exceeds $N(1/2 + p_1(n)/2)$. Setting $\mu = N(1/2 + \nu(n))$ and $z = N(1/2 + p_1(n)/2)$, we use the Chernoff bound $\Pr[X > (1+\delta)\mu] < \exp(-\mu\delta^2/4)$, which holds when $\delta < 2e - 1$. We are interested in the case $\delta = (p_1(n) - 2\nu(n))/(1+2\nu(n)) > p_1(n)/2$ and so $\Pr[X > N(1/2+p_1(n)/2)] < \exp(-N(p_1(n))^2/8)$. Again, taking, say, $N = (P_1(n))^3$, makes this quantity exponentially small in n. Thus, the probability that $\mathcal{D}$ accepts at the Step 0 of the chain is negligibly close to 1/2.

Henceforth, we say u is *good* (for the distinguisher) if it gives rise to a gap of size at least $p_A/4(\ell+1)$ in the probability that $\mathcal{D}$ will accept when $u_k = u$ in Step $k+1$ of the chain. The gap is between the probability of acceptance at Steps k and $k+1$ in the chain, and the probability space is over the choice of $u_0, \ldots, u_{k-1}$, the choice of the "public key", the randomness used by the distinguisher, and and coin flips of $\mathcal{C}$ in both cases. Clearly, the measure of good u in the unit ball is at least n^{-c_1} for some constant c_1.

6. Show how to lift a point close to zero or one families of hyperplanes, respectively, in $\mathbb{R}^n$, to a point close to $k-1$ or k families of hyperplanes, repsectively, in $\mathbb{R}^{n+\ell}$. With this step we can ensure that if we have a method of sampling points close to zero or one families of hyperplanes (we don't need to know which) in the space of lower dimension, then we can convert it into a source for points close to k or $k+1$ families, respectively, in the higher dimension.

   To do this, we choose a random n-dimensional subspace S of $\mathbb{R}^{n+\ell}$, embed $\mathbb{R}^n$ in S in the natural way, and then lift points in S to the full space, such that if the points were origially chosen from $\mathcal{H}\mathrm{Samp}_{\hat{u}}^{n,\mathcal{K}(n)} + \mathcal{N}_n(0,\rho^2)$, where $\hat{u}$ is unknown, then the lifted points are distributed according to $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_{k-1},u}^{n+\ell,\mathcal{K}(n)} + \mathcal{N}_{n+\ell}(0,\rho^2)$ in $\mathbb{R}^{n+\ell}$. Here, $u_0, \ldots, u_{k-1}$ are chosen as in the public key, but from the orthogonal complement of S, and u is the embedding of $\hat{u}$ in S. Note that the orientation of u in $\mathbb{R}^{n+\ell}$ is uniform. If instead the points were chosen from $\mathcal{N}_n(0,\mathcal{K}(n)^2) + \mathcal{N}_n(0,\rho^2)$ in $\mathbb{R}^n$, then the resulting points are distributed according to $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_{k-1}}^{n+\ell,\mathcal{K}(n)} + \mathcal{N}_{n+\ell}(0,\rho^2)$ in $\mathbb{R}^{n+\ell}$. This is done in Section 6.

   Summarizing the argument up to this point: If the cryptosystem is insecure then we can distinguish between the following two distributions of points in $\mathbb{R}^n$: $\mathcal{H}\mathrm{Samp}_v^{n,\mathcal{K}(n)} + \mathcal{N}_n(0,\rho^2)$ and $\mathcal{N}_n(0,\mathcal{K}(n)^2) + \mathcal{N}_n(0,\rho^2)$, where v is distributed as the private key when $\ell = 0$. That is, we can distinguish points chosen with a Gaussian-like distribution close to a single family of hyperplanes from points chosen with a Gaussian-like distribution without regard to any family of hyperplanes. This is done by repeatedly choosing

$u_0, \ldots, u_{k-1}$, creating public keys by lifting samples from the given distribution (either $\mathcal{H}\mathrm{Samp}_v^{n,\mathcal{K}(n)} + \mathcal{N}_n(0, \rho^2)$ or $\mathcal{N}_n(0, \mathcal{K}(n)^2) + \mathcal{N}_n(0, \rho^2)$, we don't know which) to create a "public key" for either the kth or $(k+1)$st step in the chain, and giving these points to the distinguisher $\mathcal{D}$.

7. Show that a polynomial ability to distinguish perturbed hyperplane points distributed as $\mathcal{H}\mathrm{Samp}_u^{n,\mathcal{K}(n)} + \mathcal{N}_n(0, \rho^2)$, for u chosen as in the private key when $\ell = 0$, from $\mathcal{N}_n(0, \mathcal{K}(n)^2) + \mathcal{N}_n(0, \rho^2)$ implies the ability to solve an arbitrary instance of the unique shortest vector problem. This step takes place entirely in $\mathbb{R}^n$, and deals only with a single collection of hyperplanes and in broad outline, but differing in detail, follows the proof of a similar claim in [2]. This step appears in Section 9.

8. In this final step we choose the value of $\mathcal{K}(n)$. The sizes (in bits) of the public key and the ciphertext depend on this value.

5 Sampling Lattice Points According to a Gaussian

Sampling according to a Gaussian arises both in the choice of the public key and in the proof of worst-case/average-case equivalence. The latter requires sampling from the dual of a known lattice (presented by a basis), with an unknown unique shortest vector u. In the latter case the points of the dual lie in $\mathcal{H}_u$, and we wish to chose points on hyperplanes with distribution $\mathcal{H}\mathrm{Samp}_u^{1,\mathcal{K}(n)}$, without knowing u. In addition, it arises when we lift n-dimensional points to $\mathbb{R}^{n+\ell}$ in Step 6 of the proof.

The Distribution $\mathrm{LatticeRound}_{\mathcal{B}}[\mathcal{K}(n)]$**.** For integer $n \geq 1$, let L be a lattice of arbitrary determinant presented by a basis $\mathcal{B} = (b_1, \ldots, b_n)$. The distribution $\mathrm{LatticeRound}_{\mathcal{B}}[\mathcal{K}(n)]$ is defined by the following sampling procedure:

1. Choose a point $y \in \mathbb{R}^n$ by sampling from $\mathcal{N}_n(0, \mathcal{K}(n)^2)$, without regard to the lattice.

2. Express y as a linear combination of the known basis vectors $y = \beta'_1 b_1 + \ldots + \beta'_n b_n$.

3. For $i = 1, \ldots, n$ let $\beta_i = \lfloor \beta'_i \rceil$. Let $x = \sum_{i=1}^n \beta_i b_i$.

4. Output x.

Lemma 5.1 *Let $n \geq 1$ be an integer, and assume $M > 1$, $0 < \epsilon < 1$, and R are real numbers. Let L be a lattice in $\mathbb{R}^n$ with arbitrary determinant Δ presented by a basis $\mathcal{B}$ consisting of vectors no longer than M, and assume that the sampling procedure* $\mathrm{LatticeRound}_{\mathcal{B}}[R]$ *gives the lattice point $x \in L$ with probabilty p_x. For each $x \in L$ let δ_x denote an n-dimensional "offset" for x. Let $\xi \in \mathbb{R}$ satisfy $nM \leq \xi < R^{1/4}$, and assume that $\|\delta_x\|_\infty \leq \xi$ for all $x \in L$ and $R^{1-\epsilon} > 6n^2\xi$. Then*

$$\sum_{x \in L} |p_x - g(x + \delta_x)\Delta| \leq \frac{5n\xi^2}{R^2} + \frac{6n^2\xi}{R^2} + 2ne^{-\frac{1}{2}\left(\frac{nR^\epsilon}{2}\right)^2} + e^{-\frac{1}{16}(nR^\epsilon)^2}$$

where

$$g(x) = g_R^n(x) = (R\sqrt{2\pi})^{-n} e^{-\frac{1}{2}R^{-2}\sum x_i^2}.$$

Note that if $R = 2^{\log^c m}$ for an appropriate c (as a function of ϵ) satisfies the conditions of the lemma, then we obtain yields a statistical distance negligible in m. Thus, even if $n = 1$ we obtain meaningful bounds when R is sufficiently large.

Proof: Let $K > 0$ be a real number. Our estimate on the sum in the lemma will depend on the parameter K. We get the result claimed in the lemma for $K = nR^{1+\epsilon}$, but other choices of K may yield better upper bounds. During the proof we will work with an arbitrary K and note if any other assumption on K is needed. The choice $K = nR^{1+\epsilon}$ will satisfy all of these addditional assumptions. For the moment we only assume that $K > 4\xi$ and $\frac{1}{2}R^{-2}(5n\xi^2 + 6nK\xi) < 1$ which, according the assumtpions of the lemma hold for $K = nR^{1+\epsilon}$. Let

$$\sum_{x \in L} |p_x - g(x)\Delta| = S_1 + S_2$$

where

$$S_1 = \sum_{x \in L \colon \|x\|_\infty + \|\delta_x\|_\infty < K} |p_x - g(x + \delta_x)\Delta| \; |$$

and

$$S_2 = \sum_{x \in L \colon \|x\|_\infty + \|\delta_x\|_\infty \geq K} |p_x - g(x + \delta_x)\Delta|.$$

We estimate S_1 and S_2 separately.

First we get an upper bound on S_1. Assume that x satisfies the conditions given in the definition of S_1. The probability that we select x is $p_x = \int_{x+\mathcal{P}} g(y)dy$. Let $\mathcal{P} = \mathcal{P}(b_1, \ldots, b_n)$. Since the determinant of the lattice is Δ, the continuity of the function g imples that $p_x = g(v)\Delta$ for a suitably chosen $v \in x + \mathcal{P}$. Therefore

$$\begin{aligned} |p_x - g(x)\Delta| &\leq \max_{u \in x+\mathcal{P}} |g(v)\Delta - g(u + \delta_x)\Delta| \\ &= \max_{u \in x+\mathcal{P}} \Delta |g(v) - g(u + \delta_x)| \end{aligned}$$

Assume that a point u is fixed where the maximum is attained in the last expression. Since the l_∞ diameter of $\mathcal{P}$ is at most $nM < \xi$ and $\|x\|_\infty + \|\delta_x\|_\infty < K$, we have $\sum_{i=1}^n |(u_i + (\delta_x)_i)^2 - v_i^2| = \sum_{i=1}^n |(x_i + (u_i + (\delta_x)_i) - x_i)^2 - (x_i + (v_i - x_i))^2| = \sum_{i=1}^n |((u_i + (\delta_x)_i) - x_i)^2 + 2|x_i((u_i + \delta_x)_i - x_i)| + (v_i - x_i)^2 + 2|x_i(v_i - x_i)|$ where $x = \langle x_1, ..., x_n \rangle$, $u = \langle u_1, ..., u_n \rangle$, $v = \langle v_1, ..., v_n \rangle$, and $\delta_x = \langle (\delta_x)_1, \ldots, (\delta_x)_n \rangle$.

Now, $(u_i + (\delta_x)_i - x_i)^2 \leq (|u_i - x_i| + |(\delta_x)_i|)^2 \leq (2\xi)^2$, and each $x_i < K$. Thus,

$$\sum_{i=1}^n |u_i^2 - v_i^2| \leq 5n\xi^2 + 6nK\xi$$

Therefore

$$\frac{g(u + \delta_x)}{g(v)} = e^{-\frac{1}{2}R^{-2} \sum ((u_i + (\delta_x)_i)^2 - v_i^2)} \leq e^{-\frac{1}{2}R^{-2}(5n\xi^2 + 6nK\xi)}$$

This implies that

$$\begin{aligned} \Delta |g(v) - g(u + \delta_x)| &= \Delta |g(v)(1 - \frac{g(u + \delta_x)}{g(v)})| \\ &\leq \Delta g(v) |1 - e^{-\frac{1}{2}R^{-2}(5n\xi^2 + 6nK\xi)}| \\ &= p_x |1 - e^{-\frac{1}{2}R^{-2}(5n\xi^2 + 6nK\xi)}| \end{aligned}$$

and so we have

$$|p_x - g(x+\delta_x)\Delta| \leq p_x|1 - e^{-\frac{1}{2}R^{-2}(5n\xi^2+6nK\xi)}|$$

According to our the assumptions on R and choice of K we have $\frac{1}{2}R^{-2}(5n\xi^2 + 6nK\xi) < 1$. Therefore

$$|p_x - g(x+\delta_x)\Delta| \leq p_x\frac{1}{2}R^{-2}(5n\xi^2 + 6nK\xi)$$

This is true for each x satisfying the conditions in the definition of S_1. Therefore adding the inequalities for all of these vectors x, and using that since p_x is a distribution we have $\sum p_x = 1$, we get that

$$S_1 = \sum_{x\in L:\, \|x\|_\infty+\|\delta_x\|_\infty<K} |p_x - g(x)\Delta| \leq \frac{1}{2}R^{-2}(5n\xi^2 + 6nK\xi)$$

This completes the proof of the upper bound on S_1.

We prove now an upper bound on S_2.

$$S_2 = \sum_{x\in L:\, \|x\|_\infty+\|\delta_x\|_\infty\geq K} |p_x - g(x+\delta_x)\Delta| \leq Z_1 + Z_2\Delta$$

where $Z_1 = \sum_{x\in L:\, \|x\|_\infty+\|\delta_x\|_\infty\geq K} p_x$ and $Z_2 = \sum_{x\in L:\, \|x\|_\infty+\|\delta_x\|_\infty\geq K} g(x+\delta_x)$ In the proof of the upper bound on Z_1 we will use the following well-known fact:

Fact 5.1 *If* $\varphi(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}$ *and* $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}y^2}dy = \int_{-\infty}^{x}\varphi(y)dy$ *then for every* $x > 0$ *we have*

$$\varphi(x)(\frac{1}{x} - \frac{1}{x^3}) < 1 - \Phi(x) < \varphi(x)\frac{1}{x}. \tag{2}$$

(For a proof see e.g. William Feller, An introduction to Probability Theory and its Applications, Vol. 1., Chapter VII, Lemma 2, Second Ed., John Wiley & Sons, Inc. 1961.)

The assumptions $K > 4\xi$ and $\|\delta_x\|_\infty < \xi$ together imply that if $\|x\|_\infty + \|\delta_x\|_\infty \geq K$ then $\|y\|_\infty \geq K - \|\delta_x\|_\infty - \xi \geq K - 2\xi \geq \frac{K}{2}$, where y was defined in the selection process. $Z_1 = \sum\{p_x \mid x \in L \wedge \|x\|_\infty + \|\delta_x\|_\infty \geq K\}$ is the probability of the event that $\|x\|_\infty \geq K - \xi$ and according to the last inequality this not larger than the probability of the event that $\|y\|_\infty \geq \frac{K}{2}$.

We have picked $y = \langle y_1, ..., y_n\rangle \in \mathbb{R}^n$ with distribution G. The density function of the distribution of each component y_i of y is

$$h(z) = (R\sqrt{2\pi})^{-1}e^{-\frac{1}{2}R^{-2}z^2}$$

Therefore if $J > 1$ then using inequality (2) we get

$\mathtt{prob}(y_i \geq J) < 2\int_J^\infty (R\sqrt{2\pi})^{-1}e^{-\frac{1}{2}R^{-2}z^2}dz = 2\int_{R^{-1}J}^\infty (\sqrt{2\pi})^{-1}e^{-\frac{1}{2}z^2}dz \leq 2e^{-\frac{1}{2}R^{-2}J^2}$

Consequently for all $J > 1$ we have

$$\mathtt{prob}(\|y\|_\infty \geq J) = \mathtt{prob}(\exists i \in [1,n], |y_i| \geq J,) < 2ne^{-\frac{1}{2}R^{-2}J^2}$$

Therefore, since $K > 4\xi \geq 4nM > 4n > 4$, we get $Z_1 \leq \mathtt{prob}(\|y\|_\infty \geq \frac{1}{2}K) < 2ne^{-\frac{1}{2}(\frac{K}{2R})^2}$.

Now we estimate Z_2. Let $K' = K/2$ and $K' = r_1 < r_2 < ....$ be an infinite sequence of positive real numbers with $r_{i+1} \leq 2r_i$, $i = 1, 2, ...$ and $\lim_{i\to\infty} r_i = \infty$. Assume further then in the ball B_{r_i} with radius r_i around 0 the number of lattice points is at most N_i and U_i is an upper bound on the function g in $B_{2r_i} \backslash B_{r_i}$. Then, if $H = \{x \in L \mid \|x\|_\infty > K'\}$, $r_{i+1} \leq 2r_i$ implies that $H \subseteq \bigcup_{i=1}^{\infty} B_{2r_i} \backslash B_{r_i}$ and so we have $Z_2 \leq \sum_{x \in H} g(x) \leq \sum_{i=1}^{\infty} N_i U_i$. We estimate Z_2 using this inequality with a suitable choice of the sequence r_i.

Since $K' > \xi > nM$, we have that $N_i \Delta$ is smaller than the volume of B_{3r_i}. This is true since $x \in B_{2r_i}$ implies that $x + \mathcal{P} \subseteq B_{3r_i}$. Therefore $N_i \Delta \leq \gamma_n 3^n r_i^n$, where γ_n is the volume of the unit ball. On the other hand we have

$$U_i \leq g(\langle 1, 0, \ldots, 0 \rangle) \leq (R\sqrt{2\pi})^{-n} e^{-\frac{1}{2}R^{-2}r_i^2}$$

and so

$$N_i U_i \leq \gamma_n 3^n r_i^n (R\sqrt{2\pi})^{-n} e^{-\frac{1}{2}R^{-2}r_i^2} \frac{1}{\Delta}$$

We have

$$Z_2 \Delta \leq \sum_{i=1}^{\infty} N_i U_i \Delta \leq \sum_{i=1}^{\infty} \gamma_n 3^n r_i^n (R\sqrt{2\pi})^{-n} e^{-\frac{1}{2}R^{-2}r_i^2}$$

Assume now that $r_i = iK'$. We claim that the value of the infinite series is smaller than twice its first term. Indeed the ratio of the $i+1$th and ith term is $(\frac{r_{i+1}}{r_i})^n e^{-\frac{1}{2}R^{-2}(r_{i+1}^2 - r_i^2)} = (\frac{i+1}{i})^n e^{-\frac{1}{2}R^{-2}(K')^2(2i+1)} \leq \frac{1}{2}$ provided that $K' \geq nR$ (which clearly holds for $K = nR^{1+\epsilon}$). Consequently $\sum_{i=1}^{\infty} N_i U_i \leq N_1 U_1 \sum_{i=0}^{1} 2^{-i} \leq 2N_1 U_1$

Therefore we get $Z_2 \Delta \leq 2N_1 U_1 = \gamma_n 3^n (K')^n (R\sqrt{2\pi})^{-n} e^{-\frac{1}{2}R^{-2}(K')^2} = \gamma_n (\frac{3}{\sqrt{2\pi}})^n (\frac{K'}{R})^n e^{-\frac{1}{2}(\frac{K'}{R})^2} = e^{n\log(\frac{K'}{R}) - \frac{1}{2}(\frac{K'}{R})^2} \leq e^{-\frac{1}{4}(\frac{K'}{R})^2} = e^{\frac{1}{16}(\frac{K}{R})^2}$, provided that $\frac{K'}{R} > n$ and nR^ϵ is sufficiently large.

The upper bounds on Z_1 and Z_2 imply that

$$S_2 = Z_1 + Z_2 \Delta \leq +2ne^{-\frac{1}{2}\left(\frac{nR^\epsilon}{2}\right)^2} + e^{-\frac{1}{16}(nR^\epsilon)^2}$$

and so $S_1 + S_2 \leq \frac{1}{2}R^{-2}(5n\xi^2 + 6n^2\xi) + 2ne^{-\frac{1}{2}\left(\frac{nR^\epsilon}{2}\right)^2} + e^{-\frac{1}{16}(nR^\epsilon)^2}$. ■

Let Λ be a lattice with a unique shortest vector u. Let Λ^* be presented with basis $\mathcal{B} = (b_1, \ldots, b_n)$. Since the distribution $\mathrm{LatticeRound}_{\mathcal{B}}[\mathcal{K}(n)]$ is a distribution on lattice points in Λ^*, we may ask, for any $i = 0, \pm 1, ...$, the probability that a sample from this distribution lies in H_i, the $(n-1)$-dimensional hyperplane in $\mathbb{R}^n$ containing points whose inner product with u equals i. This is addressed in the following claim.

Claim 5.1 *Let Λ be a lattice with a unique shortest vector u. Let Λ^* be presented with basis $\mathcal{B} = (b_1, \ldots, b_n)$. For integer i, let q_i denote the mass of the distribution $\mathrm{LatticeRound}_{\mathcal{B}}[\mathcal{K}(n)]$ in H_i. Let*

$$s_i = \frac{g^{(1)}_{\mathcal{K}(n)}(i/\|u\|)}{\sum_{j \in \mathbb{Z}} g^{(1)}_{\mathcal{K}(n)}(j/\|u\|)}$$

Then $\sum_{i \in \mathbb{Z}} |q_i - s_i|$ is negligible in $m > n$ provided the conditions of Lemma 5.1 are satisfied.

Proof: Note that since $\text{LatticeRound}_{\mathcal{B}}[\mathcal{K}(n)]$ returns a point in Λ^*, each value obtained belonds to some H_i. Let $\nu(m), \nu_1(m), \nu_2(m), \ldots$ be functions that are negligible in m. For $x \in \Lambda^*$ let p_x denote the mass assigned to x by $\text{LatticeRound}_{\mathcal{B}}[\mathcal{K}(n)]$, so that for $i \in \mathbb{Z}$ we have $q_i = \sum_{x \in H_i \cap \Lambda^*} p_x$.

Fact 5.2 *Let* $\Delta = |\Lambda^*|$.

$$\sum_{i \in \mathbb{Z}} \left| \left(\sum_{x \in H_i \cap \Lambda^*} p_x \right) \quad - \quad \frac{\sum_{x \in H_i \cap \Lambda^*} g^n_{\mathcal{K}(n)}(x)\Delta}{\sum_{x \in \Lambda^*} g^n_{\mathcal{K}(n)}(x)\Delta} \right| \tag{3}$$

$$\leq \quad \sum_{i \in \mathbb{Z}} \sum_{x \in H_i \cap \Lambda^*} \left| p_x \quad - \quad \frac{g^n_{\mathcal{K}(n)}(x)\Delta}{\sum_{x \in \Lambda^*} g^n_{\mathcal{K}(n)}(x)\Delta} \right| \quad \leq \quad \nu(m). \tag{4}$$

Proof: The first inequality is immediate; we prove the second. We know from Lemma 5.1 that the denominator in Equation 4 can be written as $1 \pm \nu_1(m)$. Putting everything over this common denominator we have that the quantity in Equation 4 is bounded above by

$$\sum_{i \in \mathbb{Z}} \sum_{x \in H_i \cap \Lambda^*} \left| \frac{(1 \pm \nu_1(m))p_x - g^n_{\mathcal{K}(n)}(x)\Delta}{1 \pm \nu_1(m)} \right| \tag{5}$$

$$\leq \quad \frac{1}{1 \pm \nu_1(m)} \left(\sum_{i \in \mathbb{Z}} \sum_{x \in H_i \cap \Lambda^*} \left| p_x - g^n_{\mathcal{K}(n)}(x)\Delta \right| \right) + \frac{1}{1 \pm \nu_1(m)} \sum_{i \in \mathbb{Z}} \sum_{x \in H_i \cap \Lambda^*} \nu_1(m) p_x \tag{6}$$

Of these two terms, the one on the left is negligible by Lemma 5.1, while the one on the right is negligible because the p_x sum to 1. ■

Let $L' = \Lambda^* \cap H_0$. For every $i \in \mathbb{Z}$ there is a "small" offset $\mathbf{o}_i \leq (n-1)\text{bl}(L')$, such that

$$\sum_{x \in H_i \cap \Lambda^*} g^n_{\mathcal{K}(n)}(x) \quad = \quad \sum_{y \in L'} g^1_{\mathcal{K}(n)}(i/\|u\|) g^{n-1}_{\mathcal{K}(n)}(y + \mathbf{o}_i) \tag{7}$$

$$= \quad g^1_{\mathcal{K}(n)}(i/\|u\|) \sum_{y \in L'} g^{n-1}_{\mathcal{K}(n)}(y + \mathbf{o}_i) \tag{8}$$

From this and Fact 5.2 we immediately have:

$$\sum_{i \in \mathbb{Z}} \left| \left(\sum_{x \in H_i \cap \Lambda^*} p_x \right) \quad - \quad \frac{g^1_{\mathcal{K}(n)}(i/\|u\|) \sum_{y \in L'} g^{n-1}_{\mathcal{K}(n)}(y + \mathbf{o}_i)\Delta}{\sum_{j \in \mathbb{Z}} g^1_{\mathcal{K}(n)}(j/\|u\|) \sum_{y \in L'} g^{n-1}_{\mathcal{K}(n)}(y + \mathbf{o}_j)\Delta} \right| \leq \nu(m) \tag{9}$$

Let us write $\Delta = z\Delta'$, where Δ' is the determinant of L'. As we will explain, by Lemma 5.1, with dimension $n-1$, determinant Δ', and $R = \mathcal{K}(n)$, we have

$$\frac{\sum_{y \in L'} g^{n-1}_{\mathcal{K}(n)}(y + \mathbf{o}_i)\Delta'}{\sum_{y \in L'} g^{n-1}_{\mathcal{K}(n)}(y + \mathbf{o}_j)\Delta'} \quad = \quad \frac{\sum_{y \in L'} g^{n-1}_{\mathcal{K}(n)}(y + \mathbf{o}_i)\Delta'}{(\sum_{y \in L'} g^{n-1}_{\mathcal{K}(n)}(y + \mathbf{o}_i)\Delta') \pm \nu_1(m)} \tag{10}$$

and moreover both numerator and denominator are negligibly (in m) close to 1 provided $\mathcal{K}(n)$ is sufficiently large with respect to Δ'.

The reason we can apply Lemma 5.1, even though we don't have a basis for L' and we are not using the sampling procedure LatticeRound at all, is that Equation 10 talks only about the Gaussian distribution $g^{n-1}_{\mathcal{K}(n)}$ and Δ', and not about sampling. Suppose we had (miraculously) a basis for L' of length at most M. Then we could apply Lemma 5.1 with $R = \mathcal{K}(n)$ and provided the conditions of the lemma are satisfied, in particular, that $\forall i, j \in \mathbb{Z} : \|\mathbf{o}_i - \mathbf{o}_j\|_\infty \leq \xi$, we would have that, letting r_x denote the mass assigned to $x \in L'$, for all $i \in \mathbb{Z}$, $\sum_{x \in L'} |r_x - g^{n-1}_{\mathcal{K}(n)}(x + \mathbf{o}_i)\Delta'|$ is negligible in m. Then, since the r_x sum to 1, we get that $\sum_{x \in L'} g^{n-1}_{\mathcal{K}(n)}(x + \mathbf{o}_i)\Delta'$ is negligibly close to 1.

From this and Equation 9 we get

$$\sum_{i \in \mathbb{Z}} \left| \left(\sum_{x \in H_i \cap \Lambda^*} p_x \right) - \frac{g^1_{\mathcal{K}(n)}(i/\|u\|)z}{\sum_{j \in \mathbb{Z}} g^1_{\mathcal{K}(n)}(j/\|u\|)z}(1 \pm \nu_2(m)) \right| \leq \nu_3(m). \tag{11}$$

The lemma now follows by manipulation, and the fact that the p_x sum to 1.

■

Claim 5.1 is important. It says that if we have a basis for Λ^* *with unknown unique shortest vector* u, then we can sample points in Λ^* (as we need to do in the reduction) with essentially the same distribution on choice of hyperplane as we obtain on the choice of hyperplane from the collection induced by u when choosing a public key with u part of the private key.

6 Lifting Points

We are given points y in $\mathbb{R}^n$ close to zero, or, respectively, one, collections of hyperplanes in $\mathbb{R}^n$; our goal is to lift them to $\mathbb{R}^{n+\ell}$ so that the resulting points are close to k, respectively $k+1$, collections of hyperplanes in $\mathbb{R}^{n+\ell}$.

Let S be a random n-dimensional subspace of $\mathbb{R}^{n+\ell}$, with orthonormal basis $b_1, \ldots, b_n$. Here each $b_i \in \mathbb{R}^{n+\ell}$. We embed each $y = \sum_{i=1}^n a_i \mathbf{e_i} \in \mathbb{R}^n$ into S in the natural fashion: y gets mapped to $x = \sum_{i=1}^n a_i b_i$.

Let $\hat{u} \in \mathbb{R}^n$, and let u be the embedding of $\hat{u}$ in $\mathbb{R}^{n+\ell}$. So $u \in S$, and if $y \in \mathbb{R}^n$ is at distance δ to the family of $(n-1)$-dimensional hyperplanes $\mathcal{H}_{\hat{u}}$, then its embedding $x \in \mathbb{R}^m$ is at distance δ from the family of $(n+\ell-1)$-dimensional hyperplanes $\mathcal{H}_u$.

Choose $u_0, \ldots, u_{k-1}, u_{k+1}, \ldots u_\ell$ as in the public key, only from the orthogonal complement of S in $\mathbb{R}^{n+\ell}$. Thus, u and all the u_i are mutually orthogonal, and the u_i span $S^\perp$.

We now describe how to lift a point $x \in S$ to the full space, such that if x is the embedding of y and y was chosen according to $\mathcal{H}\mathrm{Samp}^{n,R}_{\hat{u}} + \mathcal{N}_n(0, \rho^2)$, where $\hat{u}$ is unknown, then the lifted point is distributed according to $\mathcal{H}\mathrm{Samp}^{n+\ell,R}_{u_0,\ldots,u_{k-1},u} + \mathcal{N}_{n+\ell}(0, \rho^2)$ in $\mathbb{R}^{n+\ell}$. If instead x is the embedding of y and y was chosen from $\mathcal{N}_n(0, R^2) + \mathcal{N}_n(0, \rho^2)$ in $\mathbb{R}^n$, then the lifted point is distributed according to $\mathcal{H}\mathrm{Samp}^{n+\ell,R}_{u_0,\ldots,u_{k-1}} + \mathcal{N}_{n+\ell}(0, \rho^2)$ in $\mathbb{R}^{n+\ell}$.

For $j = 0, \ldots, k-1$ let i_j be an independent sample of $\mathcal{H}\mathrm{Samp}^{1,\mathcal{K}(n)}_{u_j}$. For $j = k+1, \ldots, \ell$

let i_j be an independent sample from $\mathcal{N}_1(0, \mathcal{K}(n)^2)$ and set

$$z = \sum_{j=0}^{k-1} i_j \mathbf{e_{u_j}} + \sum_{j=k+1}^{\ell} i_j \mathbf{e_{u_j}}.$$

Finally, we "complete" the perturbation of $x + z$ to the full space by adding a perturbation in $S^\perp$, choosing from $\mathcal{N}_\ell(0, \rho^2)$ in the subspace $S^\perp$ and adding the result to $x + z$ to obtain the point v.

Let $u_k = u$ and recall that u is the embedding of $\hat{u}$. The following lemma is immediate:

Lemma 6.1 *If $\hat{u}$ is chosen as in the public key with $\ell = 0$, and if (a) Y is a random variable with distribution $\mathcal{H}\mathrm{Samp}_{\hat{u}}^{n,\mathcal{K}(n)} + \mathcal{N}_n(0, \rho^2)$, the process described above yields a random variable V with distribution $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_{k-1},u}^{n+\ell,\mathcal{K}(n)} + \mathcal{N}_{n+\ell}(0, \rho^2)$. If instead (b) Y has distribution $\mathcal{N}_n(0, \mathcal{K}(n)^2) + \mathcal{N}_n(0, \rho^2)$, then the resulting V has distribution $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_{k-1}}^{n+\ell,\mathcal{K}(n)} + \mathcal{N}_{n+\ell}(0, \rho^2)$.*

Moreover, if $\hat{u}$ is chosen according to the distribution on public keys for the case $\ell = 0$ then regardless of which of the two possible distributions of Y, this procedure yields a random variable distributed exactly as one of the distributions for our distinguisher – perturbations of points close to the $k + 1$ hyperplane collections induced by $u_0, \ldots u_k$ (case (a)) or perturbations of points close to the k hyperplane collections induced by $u_0, \ldots, u_{k-1}$ (case (b)).

7 Unstructured "Dust" Foils The Adversary

In this section we prove that at Step 0 of the hybrid chain with overwhelming probability over choice of the points in D, random subset sums of elements in D have a distribution negligibly close to the uniform distribution modulo $\mathcal{P}$. Thus, at this step of the chain, with overwhelming probability, the adversary cannot distinguish between encryptions of *any* two messages m_0, m_1. It follows that, as discussed in Section 4, the distinguisher $\mathcal{D}$ accepts with probability negligibly close to 1/2 at this step in the chain. The proof holds regardless of whether or not the vertices defining $\mathcal{P}$ are close to hyperplane collections (at Step 0 of the chain they are not).

Let the points $p_1, \ldots, p_{n+\ell}$ defining a sufficiently wide parallelepiped of odd volume be according to $\mathcal{N}_{n+\ell}(0, \mathcal{K}(n)^2) + \mathcal{N}_{n+\ell}(0, \rho^2)$ Let $d_1, \ldots, d_{m'}$ be chosen according to $\mathcal{N}_{n+\ell}(0, \mathcal{K}(n)^2) + \mathcal{N}_{n+\ell}(0, \rho^2)$. We will first argue that the sum of a relatively small number of samples of $\mathcal{N}_{n+\ell}(0, \mathcal{K}(n)^2) + \mathcal{N}_{n+\ell}(0, \rho^2)$ gives a negligibly close distribution to uniform, modulo $\mathcal{P}$. Then we apply Lemma 7.2 to draw the desired conclusion about arbitrary subset sums drawn from a fixed set of samples.

Divide the d_j's into blocks of size t (to be determined later). We let b_i denote the sum of the d's in the ith block, so each b_i has distribution $\mathcal{N}_{n+\ell}(0, t(\mathcal{K}(n))^2)$.

The lattice $\mathcal{L} = L(p_1, \ldots, p_{n+\ell})$ clearly has $\lambda_n(\mathcal{L}) \leq \max_i ||p_i||$, and so we may apply results of Micciancio and Regev ([13] Lemmas 3.3 and 4.1) which, intuitively, say that if we perturb a randomly chosen lattice point in $\mathcal{L}$ by a Gaussian of parameter at least

$$\sqrt{\frac{\ln(2n(1 + 1/\epsilon))}{\pi}} \lambda_n(\mathcal{L})$$

the resulting distribution is $\epsilon/2$-uniform on the entire space. A rigorous formulation of this idea is that the resulting distribution, modded out by $\mathcal{P}$, is almost uniform on $\mathcal{P}^-$. (The Miccianco-Regev paper explains the situation in a similar way in the fifth paragraph on page 11.) From this result and the discussion in [14] regarding sampling a "random" lattice point we have that to ensure negligible distance from uniform on $\mathcal{P}^-$ it suffices to choose from a Gaussian with parameter $\sqrt{\omega(\log n)}\lambda_n(\mathcal{L})$, and then to reduce modulo the parallelepiped.

We therefore simply need to ensure that t is sufficiently large so that with overwhelming probability the length of each b_i is at least, say, $\log n \max_i ||p_i||$. With all but negligible probability the longest p_i has length at most $\log^2 n\mathrm{E}[||p_i||] = \log^2 n\sqrt{n+\ell}\sqrt{\mathcal{K}(n)}$. Assuming we are in this high probability case, taking $t = \log^6 n$ will suffice.

We now argue about arbitrary subset sums. We will use a modification of a lemma of Ajtai [1] about the distribution of subset sums in an Abelian group. The proof of Ajtai's lemma actually gives a little more then the statement of the lemma. This stronger version is formulated next. We do not give a proof of it, because the original proof provides this result without any modification.

Definition ((c_1, c_2)-uniform) 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 $\Pr[\xi \in X] \geq c_2$.

Lemma 7.1 ([1]) *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 $p_a = \Pr[a = \sum_{i=1}^k \delta_i b_i]$. Then

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

Remark 7.1 *Lemma 7.1 would suffice if we were to choose the vectors in D to be slightly longer, specifically, so that with high probability each vector in D exceeds the smoothing radius of the lattice $\mathcal{L} = L = (p_1, \ldots, p_{n+\ell})$. This may be accomplished by choosing them according to the distribution $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_\ell}^{n+\ell, \log^3 n\mathcal{K}(n)} + \mathcal{N}_{n+\ell}(0, \rho^2)$.*

In the following lemma we show that if in Lemma 7.1 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 7.2 $\exists c_1 > 0 \, \forall c_2 > 0 \, \exists c_3 > 0$ *so that: Assume A is a finite Abelian group, t and k are positive integers, $\xi_1, \ldots, \xi_k$ are independent random variables so that, for any sequence of integers $i_1 \ldots, i_t$ we have that $\eta = \sum_{j=1}^{t} \xi_{i_j}$ is (c_1, c_2)-uniform on the set A. Suppose further that $b = \langle b_1, \ldots, b_k \rangle$ is a sequence of length k whose elements are values of the random variables $\xi_1, \ldots, \xi_k$. Then with probability at least $1 - 2^{-c_3 k/t}$ we have:*

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

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

2. *$\sum_{a \in A} |p_a - |A|^{-1}| \leq |A|^{1/2} 2^{-2c_3 k/t}$*

Proof: (2) is a consequence of (1) so we prove (1) 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 $\{1, 2, \ldots, k\}$ 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 $\Pr[|J_{i,0}| \geq t \ \wedge \ |J_{i,1}| \geq t] \geq 1 - 2^{-c_4 t}$ 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 $\Lambda = \langle J_{1,0}, \ldots, J_{k',0} \rangle$ and let Φ be the set of all sequences $Y_0, ..., Y_{k'}$, $Y_i \subseteq I_i$ with the property that there are 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, writing $\eta(a) = p_a$ and $\iota(a) = \frac{1}{|A|}$, $\|\eta - \iota\|_{L_2} \leq 2^{-c_3 k/t}$, where we consider the L_2 norm according to the uniform measure μ on A, with $\mu(A) = |A|$. (We will use later that according to this measure if $\lambda(a)$, $a \in A$, is a probability distribution on A then $\|\lambda\|_{L_2} \leq 1$.) 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 $\eta = \Pr[B]\eta_B + \Pr[\neg B]\eta_{\neg B}$. Therefore

$$\|\eta - \iota\| \leq \|\Pr[B]\eta_B - \iota\| + \|\Pr[\neg B]\eta_{\neg B}\|.$$

Using that $\Pr[\neg B] \leq 2^{-c_5 k'}$ 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:

$$\begin{aligned} \|(\Pr[B] - 1)\eta_B + \eta_B - \iota\| &\leq (\Pr[B] - 1)\|\eta_B\| + \|\eta_B - \iota\| \\ &\leq 2^{-c_5 k'} + \|\eta_B - \iota\|. \end{aligned}$$

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 $\Lambda = X$. $\eta_B = \sum_{X\in\Phi} \Pr[B_X|B]\eta_X$, where η_X is the distribution of $\sum_{i=1}^{k} \delta_i b_i$ on A with the condition $\Lambda = X$.

$$\|\eta_B - \iota\| = \| \sum_{X\in\Phi} \Pr[B_X|B](\eta_X - \iota)\| \le \sum_{X\in\Phi} \Pr[B_X|B]\|\eta_X - \iota\|.$$

Since $\|\eta_X - \iota\| \le \|\eta_X\| + \|\iota\| \le 2$ 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 $\|\eta_X - \iota\| \le 2^{-c_7 k'}$. Taking into account now the randomization of b we may say that (1) 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 $\|\eta_X - \iota\| \le 2^{-c_8 k'}$. We prove this by showing that for every fixed $X \in \Phi$, $\|\eta_X - \iota\| \le 2^{-c_8 k'}$ 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, $i \in \Gamma$,

$$\|\eta_X - \iota\| \le 2^{-c_8 k'} \tag{12}$$

holds with probability 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 $\xi_i = b_i, i \in \Gamma$). We will show that (12) holds with η'_X instead of η_X which is clearly equivalent to the original (12). 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 $H_i = \sum_{j\in J_{i,1}} \xi_j - \sum_{j\in J_{i,0}} b_j$. $|W'| \ge \frac{k'}{2}$, so according to Lemma 7.1 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) $|J_{i,1}| \ge t$ 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 7.2.

■

The d_j satisfy an even stronger condition than that required by the lemma since, for the appropriate choice of t, for any $i_1, \ldots, i_t$ we have that $\sum_{j=1}^{t} d_{i_j} \bmod \mathcal{P}$ is actually super-polynomially close to uniformly distributed. With precision $2^{-\mathbf{p}}$, we have $|A| \le [\mathcal{K}(n)2^{\mathbf{p}}]^n$. To ensure distance $2^{-\mathbf{d}}$ from the uniform distribution we have:

$$\begin{aligned} |A|^{1/2} 2^{-2c_3 k/t} &< 2^{-\mathbf{d}} \\ |A|^{1/2} 2^{\mathbf{d}} &< 2^{2c_3 k/t} \\ \frac{1}{2}\log|A| + \mathbf{d} &< 2c_3 k/t \\ \frac{1}{2c_3}[\frac{n}{2}(\log \mathcal{K}(n) + \mathbf{p}) + \mathbf{d}] &< k/t \end{aligned}$$

Taking $m' = k$ and $t = \log^5 n$, it suffices to have $m' = [n(\log \mathcal{K}(n) + \mathbf{p}) + \mathbf{d}] \log^5 n$. When $\mathbf{p} = n$ and $\mathcal{K}(n) = 2^{O(n)}$ we get $m' = O(n^2 \log^5 n)$.

We have therefore proved:

Lemma 7.3 *If the points in part D of the "public key" have distribution $\mathcal{N}_{n+\ell}(0, \mathcal{K}(n)^2) + \mathcal{N}_{n+\ell}(0, \rho^2)$ (i.e., not close to any hyperplane collection) then, with overwhelming probability over the choice of D, the distribution on random subset sums of elements of D modulo $\mathcal{P}$ is negligibly (in n) statistically close to uniform on $\mathcal{P}^-$.*

Corollary 7.1 *At Step 0 of the chain, with overwhelming probability over choice of $D \in [\mathcal{N}_{n+\ell}(0, \mathcal{K}(n)^2) + \mathcal{N}_{n+\ell}(0, \rho^2)]^{m'}$, the distribution $\sum_{j=1}^{m'} \delta_j d_j \bmod \mathcal{P}$, where each d_j is chosen independently and uniformly from $\{0, 1\}$, is neglibly in n close to uniform on $\mathcal{P} \cap 2^{-\mathbf{p}}\mathbb{Z}$.*

Since $\mathcal{P}$ is chosen to have odd volume, the same is true for the distribution $\sum_{j=1}^{m'} 2\delta_j d_j \bmod \mathcal{P}$. Thus, at Step 0 of the chain any encoding $\sum_i b_i v_i$ of a message $b_0 b_1 \ldots, b_\ell$ is obliterated by the addition of the second term in the ciphertext $[\sum_{i=0}^{\ell} b_i v_i + \sum_{j=1}^{m'} 2\delta_j d_j] \bmod \mathcal{P}$; "ciphertexts" of *any* two messages have essentially the same distribution – the uniform distribution modulo $\mathcal{P}$.

8 Creating Step k or $k+1$ of the Chain

Let ξ be a distribution on points in $\mathbb{R}^n$. For now, we may think of ξ as either $\mathcal{H}\mathrm{Samp}_u^{n,\mathcal{K}(n)} + \mathcal{N}_n(0, \rho^2)$ or $\mathcal{N}_n(0, \mathcal{K}(n)^2) + \mathcal{N}_n(0, \rho^2)$.

Construction of a "Public Key" from ξ. We first argue that a random square binary matrix is invertible modulo 2 with constant probability. We need this in the construction of the cryptosystem in order to be able to decrypt correctly. It is also used in the proof of security. The proof of Lemma 8.1 below is based on the well-known product formula about the number of invertible matrices modulo 2. To make the paper more self-contained we prove this formula as well.

Lemma 8.1 *For any positive integer n, a random $n \times n$ binary matrix A is invertible modulo 2 with constant probability.*

Proof: The set $\{0, 1\}^n$ is a group under addition (componentwise, modulo 2). Any k linearly independent vectors in $\{0, 1\}^n$ generate a subgroup of $\{0, 1\}^n$ of size 2^k.

The probability that the first column of A is 0^n is $1/2^n$. In general, the probability that column $k+1$ is in the subgroup generated by the first k columns, given that the first k columns are linearly independent, is $\frac{2^k}{2^n}$.

If the columns are chosen one at a time, the probability that all n are linearly independent is

$$\left(1 - \frac{2^0}{2^n}\right)\left(1 - \frac{2^1}{2^n}\right)\ldots\left(1 - \frac{2^{n-1}}{2^n}\right) \quad = \quad \Pi_{i=0}^{n-1}\left(1 - \frac{2^i}{2^n}\right)$$

To prove nonzero convergence we show the inverse, $\Pi_{i=0}^{n-1}\left(\frac{2^n}{2^n - 2^i}\right)$, converges.

Calculation: $\frac{2^n}{2^n-2^i} < 1 + 2\frac{2^i}{2^n}$

$$\begin{aligned} \Pi_{i=0}^{n-1}\left(1+\frac{2^i}{2^n}\right) &< \Pi_{i=1}^{n}\left(1+\frac{2^i}{2^n}\right) & (13)\\ &= 3\Pi_{i=1}^{n-2}\left(1+\frac{2^i}{2^n}\right) & (14) \end{aligned}$$

$$\begin{aligned} \Pi_{i=1}^{n-2}\left(1+\frac{2^i}{2^n}\right) &= 1+\sum_{i=1}^{n-2}\frac{2^i}{2^n}+\sum_{\substack{i<j\\ i,j=1}}^{n-2}\frac{2^{i+j}}{2^{2n}}\dots & (15)\\ &< 1+\sum_{i=1}^{n-2}\frac{2^i}{2^n}+\sum_{i,j=1}^{n-2}\frac{2^{i+j}}{2^{2n}}+\ldots+\sum_{i_1,\ldots,i_{n-2}=1}^{n-2}\frac{2^{i_1+\ldots+i_{n-2}}}{2^{n(n-2)}} & (16) \end{aligned}$$

The sum in Equation 16 is bounded above by the infinite series $1+a_1+a_2+\ldots$ in which

$$\begin{aligned} a_k &= \sum_{i_1,\ldots,i_k=1}^{n-2}\frac{2^{i_1+\ldots i_k}}{2^{kn}} & (17)\\ \frac{a_{k+1}}{a_k} &= \frac{\sum_{i_1,\ldots,i_{k+1}=1}^{n-2}\frac{2^{i_1+\ldots i_{k+1}}}{2^{(k+1)n}}}{\sum_{i_1,\ldots,i_k=1}^{n-2}\frac{2^{i_1+\ldots i_k}}{2^{kn}}} & (18)\\ &= \frac{\frac{1}{2^{n(k+1)}}}{\frac{1}{2^{nk}}}\frac{\sum_{i_1,\ldots,i_k}2^{i_1+\ldots+i_k}\sum_{i_{k+1}}2^{i_{k+1}}}{\sum_{i_1,\ldots,i_k}2^{i_1+\ldots+i_k}} & (19)\\ &= \frac{1}{2^n}\sum_{i_{k+1}=1}^{n-2}2^{i_{k+1}} < \frac{1}{2} & (20) \end{aligned}$$

■

Sampling ξ gives a Gaussian distribution on points in R^n that are close to either zero or one family of hyperplanes. We use the lifting technique described Section 6 to lift these samples of ξ to obtain points in $R^{n+\ell}$ close to either k or $k+1$ families of hyperplanes, respectively. In this way we build the "public key" for $\mathcal{C}$.

We first randomize the parallelepiped $\mathcal{P}$, carrying out the following steps until a wide parallelepiped is obtained:

1. Choose $u_0,\ldots,u_{k-1}$.

2. Obtain $n+\ell$ samples of ξ. These are points in R^n.

3. Lift these points to $R^{n+\ell}$, so that the lifted points $p_1,\ldots,p_{n+\ell}$ are close to $\mathcal{H}_{u_0,\ldots,u_{k-1}}$ (in addition to possibly being close to $\mathcal{H}_u$, where u is a unique shortest vector of Λ). More precisely, the resulting distribution is either $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_{k-1}}^{n+\ell,\mathcal{K}(n)}+\mathcal{N}_{n+\ell}(0,\rho^2)$ or $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_k}^{n+\ell,\mathcal{K}(n)}+\mathcal{N}_{n+\ell}(0,\rho^2)$.

4. Test if the resulting parallelepiped is sufficiently wide. If not, repeat the entire procedure. If it is sufficiently wide, follow the procedure in the definition of the public key to ensure that $2^{-n}\mathbb{Z}^{n+\ell} \cap \mathcal{P}^-$ contains an odd number of points.

Lemma 8.2 *Each iteration of the procedure for obtaining a wide parallelepiped succeeds with constant probability.*

Proof: We argue that with constant probability for all $1 \leq i \leq n+\ell$ the distance from p_i to the span of $(p_1, \ldots, p_{i-1}, p_{i+1}, \ldots p_{n+\ell})$ is least roughly $2/((n+\ell)e^{\sqrt{2/\pi}})$. For each i, let $d(i)$ denote the distance $\mathrm{dist}(p_i, \mathrm{span}(p_1, \ldots, p_{i-1}, p_{i+1}, \ldots p_{n+\ell}))$. Since p_i is Gaussian, $d(i)$ is Gaussian.

$$\begin{aligned}
\Pr[d(i) < \sigma/(n+\ell)] &< \int_{-\mathcal{K}(n)/(n+\ell)}^{\mathcal{K}(n)/(n+\ell)} g^1_{\mathcal{K}(n)}(0)dx \\
&= \frac{2\mathcal{K}(n)}{n+\ell}\frac{1}{\mathcal{K}(n)\sqrt{2\pi}} \\
&= \frac{2}{(n+\ell)\sqrt{2\pi}} .
\end{aligned}$$

If the width of the parallelepiped is smaller than $\sigma/(n+\ell)$ then at least one of the distances d_i is also smaller than $\sigma/(n+\ell)$. Therefore using our upper bound on $\Pr[d(i) < \sigma/(n+\ell)]$ we get that the probability that the width is smaller than $\sigma/(n+\ell)$ is at most $n\frac{2}{(n+\ell)\sqrt{2\pi}} \leq \sqrt{\frac{2}{\pi}} < 1$. ■

Once we have the parallelepiped, we generate an additional $n+\ell+m'$ samples of ξ for the remainder of the public key, the sets V and D.

9 Worst-Case/Average-Case Equivalence

In this section we show that the ability to distinguish points close to a hyperplane collection $\mathcal{H}_u$ from points chosen without regard to the collection of hyperplanes implies the ability to solve the $\tilde{O}(n^2)$-unique shortest vector problem. Throughout this section we denote the distinguisher by $\mathcal{A}$ and assume the distinguishing gap is $n^{-\mathrm{g}}$ for some $\mathrm{g} > 0$.

A similar argument was made in [2], where inputs to the distinguisher were created either by choosing points uniformly within a large bounding box or by perturbing lattice points. If the lattice has an n^c-unique shortest vector then the points of the dual lattice are contained in a collection of hyperplanes (whose inner product with the unique shortest vector is an integer), and if the perturbation is suitable then the projection of the perturbed points onto such a hyperplane is close to uniform – essentially all structure within the hyperplane is erased by the perturbation. Since then, Regev has shown that a Gaussian with relatively small parameter yields a suitable perturbation, and as a result if the lattice has an $n^{1.5}$-unique shortest vector, and if the lattice is scaled so that its unique shortest vector has length in $[1/2, 1]$, permits us to perturb points in the dual with distribution $\mathcal{N}_n(0, 1/n^2)$ erases the structure of the dual with the hyperplanes.

Specifically, [14] Lemma 3.11 shows that if all the non-zero vectors in a lattice have length more than $\sqrt{n}$ then the distribution obtained by choosing a "random" point in the dual lattice and adding a Gaussian of parameter $1/\sqrt{2\pi}$ is exponentially close to the uniform distribution. We need to re-scale the lattice to the case in which the non-zero vectors have length more than $n^{1.5}$. This "shrinks" the basis vectors in the dual by a factor of n, so perturbing with a Gaussian of parameter $\frac{1}{\sqrt{2\pi}n}$ suffices. This parameter, needed for the proof of security, is much larger than the upper bound on ρ needed for correct decoding. This translates to a worse approximation factor: $n^{\tilde{O}(5)}$. In Section 10, we will modify the cryptosystem to obtain the approximation factor $\tilde{O}(n^2)$.

Alternatively, by Lemma 3.2 of [13], the *smoothing parameter* $\eta_{2^{-n}}(\Lambda^*)$ for Λ^* is at most $\sqrt{n}/\lambda_1(\Lambda)$. If we need $\rho < h(n)$ for correct decoding, then this serves as a smoothing factor for a lattice whose dual has no vectors of length less than $f(n) = \sqrt{n}/h(n)$. When $h(n) = \tilde{O}(n^{-4.5})$ this yields an approximation factor of $\tilde{O}(n^5)$. Later, we will get to $h(n) = \tilde{O}(n^{-1.5})$, for an approximation factor of $\tilde{O}(n^2)$.

9.0.1 Structure of the Proof

To find a unique shortest vector $u \in L$ it suffices to find H_0, the $(n-1)$-dimensional hyperplane orthogonal to u and passing through the origin. Given a basis for L we apply a randomly chosen transformation, in which the lattice is randomly spun (achieved by multiplying the basis by a random unimodular transformation) and re-sized (achieved by multiplying the basis by a random scalar). The transformation is run many times and it is argued that with high probability at least one of these results in a unique shortest vector in the transformed lattice that "looks like" a private key when $\ell = 0$: its length is in $[1/2, 1]$ and its orientation is completely random. For the remainder of the proof we let Λ be the result of such a "good" transformation. We let u denote a unique shortest vector of Λ, and Λ^* the dual of Λ. Our goal is to find $H_0 = \{v \in \mathbb{R}^n \mid u \cdot v = 0\}$.

The heart of the proof uses the distinguisher $\mathcal{A}$ to create a test, described below, of whether a point v is within a relatively narrow strip of $\mathbb{R}^n$ containing H_0. This in turn permits us to "grow" long vectors close to H_0 and use these long vectors to approximate H_0. For technical reasons we will choose $\mathcal{K}(n)$ to ensure that "long" is roughly $(\mathcal{K}(n))^{1/4}$.

In more detail, we need to obtain n mutually orthogonal vectors $w_1, \ldots, w_n$ of length at least N, all within distance $2d = 2/||u||$ of H_0. Note that the assumption that the transformation is good implies that $d \in [1, 2]$. Suppose we have already obtained $w_1, \ldots, w_{i-1}$. We will search for w_i in the $(n-i+1)$-dimensional subspace S^{n-i+1} orthogonal to $\mathrm{span}(w_1, \ldots, w_{i-1})$, such that w_i is close to $H_0 \cap S^{n-i+1}$. We now describe the general step of searching for the next starting point in the construction of w_i, assuming that we have so far found $y \in \mathrm{span}(w_1, \ldots, w_{i-1})^\perp$. Choose a random $x \in \mathrm{span}(w_1, \ldots, w_{i-1})^\perp$ according to the distribution $\mathcal{N}_{n-i}(0, d^2)$ such that $||y + x|| > ||y||$, and consider $z = y + x$ (we may take $d = 1$ for this purpose). We will run an experiment that distinguishes the two cases: (1) $|z \cdot u|$ is "sufficiently" less than 1 and (2) $|z \cdot u|$ is "not sufficiently" less than 1. There may be a gap between "sufficiently" and "not sufficiently". Let $\hat{c} > 0$ be a fixed constant.

The experiment works by examining each of $y + x/n^{\hat{c}}, y + 2x/n^{\hat{c}}, \ldots, y + x = z$ in turn. Letting v denote the vector under examination, sample from a Gaussian-like distribution defined by v; we call the distribution δ_v; it is defined later. The important property of δ_v

is: if v is sufficiently close to H_0 then δ_v contains perturbed hyperplane points, that is, δ_v is very close to $\mathcal{H}\mathrm{Samp}_u^{n,\mathcal{K}(n)} + \mathcal{N}_n(0,\rho^2)$, while if v is close to some $H_{i\neq 0}$ then δ_v is close to $\mathcal{N}_n(0,\mathcal{K}(n)^2) + \mathcal{N}_n(0,\rho^2)$. We may therefore use $\mathcal{A}$ to determine if v is close to some $H_{i\neq 0}$. We say the test is *positive* if in indicates that v is close to some $H_{i\neq 0}$.

If none of $y + x/n^{\hat{c}}, y + 2x/n^{\hat{c}}, \ldots, y + x = z$ tests positive, ie, close to some $H_{i\neq 0}$, and $||z|| \geq N$, then set $w_i = z$. If no test is positive but $||z|| < N$, then we set the next starting point to be $2z$.

9.0.2 Proof of Equivalence: Details

Lemma 9.1 *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 $\theta \in \mathrm{I\!R}$ and ν_1 is an orthogonal linear transformation of $\mathrm{I\!R}^n$;*
2. *If X is a Lebesgue measurable subset of the unit ball of $\mathrm{I\!R}^n$ whose density in it is at least n^{-c_1} and $v \in \mathrm{I\!R}^n$ with $2^{-n^2} \leq \|v\| \leq 2^{n^2}$, then $\Pr[\nu v \in X] > n^{-c_2}$.*

For us, X is the subset of the unit ball consisting of *good* (for the adversary) vectors u (as defined in the proof outline in Section 4).

Proof: The proof of this lemma uses the following facts about orthogonal linear transformations. The set of all orthogonal linear transformations of $\mathrm{I\!R}^n$ is a compact topological group under the multiplication of linear transformations and the usual topology of linear transformations (induced by, e.g., any fixed matrix representation). There is a unique proabiltiy measure on this group (defined on all Borel sets) with 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 $\mathrm{I\!R}^n$ according to this distribution. We will use the following property of μ: if $v \in \mathrm{I\!R}^n$, $\|v\| = 1$, 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 $\mu e_1, \ldots, \mu e_n$. After $\mu e_1, \ldots, \mu e_i$ has been selected, μe_{i+1} is chosen with uniform distribution from the set of all unit vectors orthogonal to $\mu e_1, \ldots, \mu e_i$.

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 $\mathcal{B}^n(1)$, and let $\beta = \|w\|$. Let γ be the random variable with takes the value $(1+\frac{1}{n})^i$ with probability $\frac{1}{2n^4+1}$ for $i = n^4, \ldots, -1, 0, 1, \ldots, n^4$.

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 $v \in \mathrm{I\!R}^n$ is fixed with $2^{-n^2} \leq \|v\| \leq 2^{n^2}$. 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 $1 \leq \gamma_0\|v\| \leq (1+\frac{1}{n})$.

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, $P(\nu v \in X | \gamma = \gamma_0) \le \frac{1}{3}n^{-c_1}$ and so $P(\nu v \in X) \le \frac{1}{3}n^{-c_1}\frac{1}{2n^2+1}$. ■

Let u and Λ^* be as described above, and assume $u \in X$. The distribution δ_v will require sampling from Λ^*. Let $\mathcal{B} = (b_1, \ldots, b_n)$ be a known basis for Λ^*. We may assume without loss of generality that $\mathcal{B}$ is reduced, for example, using the LLL algorithm, so that for $i = 1, \ldots, n$:

$$\|b_i\| \le 2^n \mathrm{bl}(\Lambda^*) \le 2^n(Cn^{1.5}/\|u\|) \tag{21}$$

where C is a universal constant [5].

Definition (the distribution δ_v). The distribution δ_v is defined by the following sampling procedure, in which x, α, and r are chosen independently. Sample x from $\mathrm{LatticeRound}_{\mathcal{B}}[\mathcal{K}(n)]$. Choose $\alpha \in_R [0,1]$. Choose r according to $\mathcal{N}_n(0,\rho^2)$. Set $\delta_v = x + \alpha v + r$.

Intuitively, if $v \in H_0$ (or even if v is only close to H_0), then αv lies in (or is close to) H_0, and so δ_v will consist of points close to hyperplanes in $\mathcal{H}_u$. On the other hand, suppose $v \in H_{i\neq 0}$ (or is close to a coset $H_{i\neq 0}$). In this case $\mathrm{dist}(\alpha v, \mathcal{H}_u)$ is (close to) uniform in $[0, 1/||u||]$. This is made formal in Lemma 9.3 below.

Lemma 9.2

1. *The distributions* $\mathcal{H}\mathrm{Samp}_u^{n,\mathcal{K}(n)} + \mathcal{N}_n(0,\rho^2)$ *and* $\mathcal{H}\mathrm{Samp}_u^{n,\mathcal{K}(n)} + \mathbf{e_u}\mathcal{N}_1(0,\rho^2)$ *have statistical distance negligibly close in* n.

2. *The distributions* $\mathcal{H}\mathrm{Samp}_u^{n,\mathcal{K}(n)} + \mathcal{N}_n(0,\rho^2)$ *and* $\mathrm{LatticeRound}_{\mathcal{B}}[\mathcal{K}(n)] + \mathcal{N}_n(0,\rho^2)$ *are negligibly close in* n.

Proof: Part 1 says that within a hyperplane there is essentially no difference between sampling from $\mathcal{N}_{n-1}(0,\mathcal{K}(n)^2)$ and sampling from $\mathcal{N}_{n-1}(0,\mathcal{K}(n)^2) + \mathcal{N}_{n-1}(0,\rho^2)$. The claim is immediate from the fact that for all $\delta \in \mathbb{R}^{n-1}$ with $\|\delta\|$ sufficiently small with respect to $\mathcal{K}(n)$, $\mathcal{N}_{n-1}(0,\mathcal{K}(n)^2)$ and $\mathcal{N}_{n-1}(\delta,\mathcal{K}(n)^2)$ are negligibly close (see, eg, the proof of Lemma 9.3 below). In our case, δ is drawn from $\mathcal{N}_{n-1}(0,\rho^2)$, and so indeed, with high probability it has very small length with respect to $\mathcal{K}(n)$.

Part 2 says that rounding to a hyperplane and then perturbing yields a distribution very close to rounding to a lattice point and then perturbing. Let C be a minimal length basis of $\Lambda^* \cap H_0$ (we don't have to be able to find it).

We have the following thought experiment:

1. Think of $\mathcal{N}_n(0,\rho^2)$ as $\mathcal{N}_1(0,\rho^2) + \mathcal{N}_{n-1}(0,\rho^2)$, it being understood that the one-dimensional normal is parallel to u and the $(n-1)$-dimensional normal is in H_0.

2. Replace $\mathcal{N}_{n-1}(0,\rho^2)$ with $\mathcal{N}_{n-1}(0,\rho^2) \bmod \mathcal{P}(C)$. This will introduce an "offset" of expected magnitude that with all but negligible probability is bounded by $\log^2 n\sqrt{n}\,\rho$ and leaves us with a distribution that is negligibly close to uniform on $\mathcal{P}(C)$.

3. Replace $\mathcal{H}\mathrm{Samp}_u^{n,\mathcal{K}(n)}$ with the distribution obtained by choosing a hyperplane according to $\mathcal{H}\mathrm{Samp}_u^{1,\mathcal{K}(n)}$ (this part is the same as in $\mathcal{H}\mathrm{Samp}_u^{n,\mathcal{K}(n)}$) but then rounding the $\mathcal{N}_{n-1}(0,\mathcal{K}(n)^2)$ selection to C, as in $\mathrm{LatticeRound}_C[\mathcal{K}(n)]$. Applying Claim 5.1 to the choice of hyperplane and Lemma 5.1 to the choice of lattice point in $H_0 \cap \Lambda^*$, and using the fact that, for every $i \in \mathbb{Z}$, $H_i \cap \Lambda^*$ is a small translation of $H_0 \cap \Lambda^*$, this gives us a distribution q_x on lattice points where
$$\sum_{x\in\Lambda^*}\left|q_x - \frac{g^n(x)\|u\|^{-1}|C|}{\sum_{z\in\Lambda^*} g^n(z)\|u\|^{-1}|C|}\right| \in \nu(n).$$

4. By Lemma 5.1, $\mathrm{LatticeRound}_{\mathcal{B}}[\mathcal{K}(n)]$ yields a distribution p_x where
$$\sum_{x\in\Lambda^*}\left|p_x - \frac{g^n(x)|\mathcal{B}|}{\sum_{z\in\Lambda^*} g^n(z)|\mathcal{B}|}\right| \in \nu(n).$$

Finally, by Lemma 5.1, we can ignore the offsets introduced in Step 2, as well as the fact that $\mathcal{N}_{n-1}(0,\mathcal{K}(n)^2) \bmod \mathcal{P}(C) + \mathcal{N}_{n-1}(0,\rho^2) \bmod \mathcal{P}(C)$ may not lie in $\mathcal{P}(C)$, adding slightly to the offsets in the analysis of $\mathcal{H}\mathrm{Samp}_u^{n,\mathcal{K}(n)}$, because with overwhelming probability these offsets are very, very small compared to $\mathcal{K}(n)^{1/4}$ provided $\mathcal{K}(n)$ is superpolynomial in n. ■

Lemma 9.3 *For all $c > 0$ there exist $c_1 > 0$ and $\rho_c \geq \rho/(n+\ell)^{c_1}$ such that*

1. *If $\mathrm{dist}(v, H_0) \leq \rho_c$, then the statistical distance between the distributions δ_v and $\mathcal{H}\mathrm{Samp}_u^{n,\mathcal{K}(n)} + \mathcal{N}_n(0,\rho^2)$ is at most n^{-c}.*

2. *If $\mathrm{dist}(v, H_{i\neq 0}) \leq \rho_c$, then the distributions δ_v and $\mathcal{N}_n(0,\mathcal{K}(n)^2) + \mathcal{N}_n(0,\rho^2)$ have statistical distance at most n^{-c}.*

It is an immediate consequence of the lemma that if $n^{-c} \ll n^{-g}$ then, since $\mathcal{A}$ distinguishes $\mathcal{H}\mathrm{Samp}_u^{n,\mathcal{K}(n)} + \mathcal{N}_n(0,\rho^2)$ from $\mathcal{N}_n(0,\mathcal{K}(n)^2) + \mathcal{N}_n(0,\rho^2)$ with gap at least n^{-g}, $\mathcal{A}$ distinguishes with polynomial gap points close to (within distance ρ/n^{c_1} of) H_0 from points close to some $H_{i\neq 0}$.

Proof: We argued in Lemma 9.2 that $\mathcal{H}\mathrm{Samp}_u^{n,\mathcal{K}(n)} + \mathcal{N}_n(0,\rho^2)$ is negligibly close to $\mathrm{LatticeRound}_{\mathcal{B}}[\mathcal{K}(n)] + \mathcal{N}_n(0,\rho^2)$.

Our concern for the first part is therefore the effect of adding αv. First, since $\|\alpha v\| \leq \mathcal{K}(n)^{1/4}$ we have by arguments analogous to those in the proof of Lemma 5.1 that $\mathcal{N}_n(\alpha v, \mathcal{K}(n)^2)$ is negligibly close to $\mathcal{N}_n(0,\mathcal{K}(n)^2)$. Thus, if v were actually in H_0 we would be done by applying Lemma 5.1 with dimension $n-1$ (describing sampling in the chosen hyperplane – recall that $\mathcal{H}\mathrm{Samp}_u^{\cdot}$ first selects a hyperplane in $\mathcal{H}_u$ and then chooses uniformly from an $(n-1)$-dimensional Gaussian in the chosen hyperplane). The effect of adding αv is in distance from H_0. We must therefore show that if $\mathrm{dist}(\alpha v, H_0)$ is sufficiently small with respect to ρ, then the difference between $\mathcal{N}_1(\mathrm{dist}(\alpha v, H_0), \rho^2)$ and $\mathcal{N}_1(0,\rho^2)$ is at most n^{-c}. This follows from arguments similar to, but simpler than, the proof of Lemma 5.1, included here for completeness.

For real numbers δ and K, let $|\delta| < \xi\rho$ and $0 \leq \xi < K/2$. Let $g(x) = g_\rho^1(x) = \frac{1}{\rho\sqrt{2\pi}} e^{-x^2/2\rho^2}$. Assume that $\frac{1}{2}(2K\xi + \xi^2) < 1$ (later we will ensure this). We are interested in bounding

$$\begin{aligned} & \int_{x \in \mathbb{R}} |g(x+\delta) - g(x)| dx & (22) \\ = & \int_{|x|<K\rho} |g(x+\delta) - g(x)| dx + \int_{|x| \geq K\rho} |g(x+\delta) - g(x)| dx & (23) \end{aligned}$$

Working first with the term on the left, we have

$$\begin{aligned} \int_{|x|<K\rho} |g(x+\delta) - g(x)| dx &= \int_{|x|<K\rho} g(x) |1 - \frac{g(x+\delta)}{g(x)}| dx \\ &= \int_{|x|<K\rho} g(x) |1 - e^{-\frac{1}{2}(2x\delta + \delta^2)/\rho^2}| dx \\ &\leq \int_{|x|<K\rho} g(x) |1 - e^{-\frac{1}{2}(2x\xi\rho + \xi^2\rho^2)/\rho^2}| dx \\ &\leq \int_{|x|<K\rho} g(x) |1 - e^{-\frac{1}{2}(2K\rho\xi\rho + \xi^2\rho^2)/\rho^2}| dx \\ &\leq \int_{|x|<K\rho} g(x) |1 - e^{-\frac{1}{2}(2K\xi + \xi^2)}| dx \\ &\leq \int_{|x|<K\rho} g(x) \frac{1}{2}(2K\xi + \xi^2) dx \\ &\leq K\xi + \frac{1}{2}\xi^2 \end{aligned}$$

Turning now to the term on the right in Equation 22, we have

$$\int_{|x| \geq K\rho} |g(x+\delta) - g(x)| dx \quad \leq \quad \int_{|x| \geq K\rho} g(x+\delta) dx + \int_{|x| \geq K\rho} g(x) dx$$

Since $|x| \geq K\rho$ and $\delta < (K/2)\rho$ we have $|x+\delta| \geq (K/2)\rho$. Thus

$$\begin{aligned} & \int_{|x| \geq K\rho} g(x+\delta) dx + \int_{|x| \geq K\rho} g(x) dx \\ \leq & \; 2 \int_{|x| \geq \frac{K}{2}\rho} g(x+\delta) dx \end{aligned}$$

Translating to the standard normal, we get $2 \int_{|x| > K/2} g_1^1(x) dx$. Using that $1 - \Phi(K/2) \leq \frac{1}{K/2}((\sqrt{2\pi})^{-1} e^{-\frac{1}{2}(K/2)^2}$, this is bounded by $\frac{4}{K}((\sqrt{2\pi})^{-1} e^{-\frac{1}{2}(K/2)^2}$. We can make this term negligible in n by choosing $K = \log^{1+\epsilon} n$. So we need that

$$K\xi + \xi^2 + \nu(n) < n^{-c}$$

which, for sufficiently large n, holds for $\xi < n^{-(c+1)}$. Thus, as long as $\mathrm{dist}(v, H_0) \leq n^{-(c+1)}$ the first part of the lemma holds. When $\rho \leq n^{-1}$ and $\ell = n$ we can take $c_1 = c+1$ to satisfy the first condition.

Similarly, for the second part of the lemma our only concern is the effect of αv on the distance of the samples of δ_v to the collection of hyperplanes. Let $H_{i \neq 0}$ be the nearest hyperplane to v in the collection. Then the distance of the distribution αv (induced by choosing $\alpha \in_R [0,1]$) and the uniform distribution on $[0, 1/||u||]$ is at most $2\text{dist}(v, \mathcal{H}_u)/d \leq 4\,\text{dist}(v, \mathcal{H}_u)$. ■

Lemma 9.4 *Let $\mathcal{A}$ distinguish any two distributions, say, ξ_0 and ξ_1, with gap $n^{-\mathrm{g}}$. Then for any $c > 0$ in polynomial time, using $\mathcal{A}$ as an oracle, we can distinguish ξ_0 and ξ_1 with gap at least $1 - n^{-c}$.*

Proof: Standard. ■

Lemma 9.5 *Let σ denote the points in $\mathbb{R}^n$ within distance $d = 1/\|u\|$ of H_0. There exist $c_3, c_4, c_5 > 0$ such that with probability $1 - n^{-c_4}$, if the "growing" procedure described above is run from a starting point y within distance $2d$ of H_0, then within n^{-c_5} iterations it produces an output that, with probability $1 - n^{-c_3}$, is in σ.*

Proof: By Lemmas 9.3 and 9.4, for any $c, c' > 0$ we can ensure that if y and z are within n^{-c} of H_0, then with probability $1 - n^{-c'} n^{\hat{c}}$ no point tests positive and hence z will not be discarded. Moreover, if $z = y + x$ is outside σ (whether or not y is outside of σ) then since the component of x in the direction perpendicular to H_0 is, with overwhelming probability, short but not very short, with polynomial probability at least one of the v's will test positive.

Finally, we need to argue that for any c, with polynomial probability $z = y + x$ is within n^{-c} of H_0. Let y be any point within distance $2d \in [2, 4]$ of H_0. Since x is distributed according to $\mathcal{N}_n(0,1)$ we can focus on its component perpendicular to H_0. It is immediate from the properties of the normal distribution that with polynomial probability the magnitude of x in this component lies within $[\text{dist}(y, H_0) - n^{-c}, \text{dist}(y, H_0) + n^{-c}]$, and half of that mass is in the direction of H_0. ■

Corollary 9.1 *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 9.1, we can find $n - 1$ mutually orthogonal long vectors $y_1, \ldots, y_{n-1}$ close to H_0. Let $\hat{H}$denote the approximation to H_0 defined by the vectors $\{y_1, \ldots, y_{n-1}\}$. We measure the quality of the approximation by finding the distance between the unit vectors orthogonal, respectively, to H_0 and $\hat{H}$. Given $n-1$ vectors $y_1, \ldots y_{n-1} \in \mathbb{R}^n$, define a generalization of the cross product $\bigotimes(y_1, \ldots, y_{n-1})$ 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 with rows $e_i, y_1, \ldots y_{n-1}$. The key point is that for any $v \in \mathbb{R}^n$,

$$v \cdot \bigotimes(y_1, \ldots, y_{n-1}) = \begin{vmatrix} v_1 & v_2 & \cdots & v_n \\ y_{11} & y_{12} & \cdots & y_{1n} \\ \vdots & & & \vdots \\ y_{n-1,1} & y_{n-1,2} & \cdots & y_{n-1,n} \end{vmatrix}.$$

In particular, for $1 \leq i \leq n-1$, $y_i \cdot \bigotimes(y_1, \ldots, y_{n-1}) = 0$. Let x be a unit vector in the direction of $\bigotimes(y_1, \ldots, y_{n-1})$. Then

$$\begin{aligned} x \cdot \bigotimes(y_1, \ldots, y_{n-1}) &= \|x\| \, \| \bigotimes(y_1, \ldots, y_{n-1})\| \cos(0) \\ &= \| \bigotimes(y_1, \ldots, y_{n-1})\|. \end{aligned}$$

But $x \cdot \bigotimes(y_1, \ldots, y_{n-1}) = \det(x, y_1, \ldots y_{n-1})$ which is the volume of the parallelepiped $\mathcal{P}(x, y_1, \ldots y_{n-1})$, which, since $\|x\| = 1$, is the volume of the parallelepiped $\mathcal{P}(y_1, \ldots, y_{n-1})$. So, since $\|x\| = 1$, $\| \bigotimes(y_1, \ldots y_{n-1})\|$ equals the volume of the parallelepiped $\mathcal{P}(y_1, \ldots, y_{n-1})$. Finally, $\bigotimes(y_1, \ldots y_{n-1})$ has positive orientiation: $\det(\bigotimes(y_1, \ldots y_{n-1}), y_1, \ldots y_{n-1}) = \bigotimes(y_1, \ldots, y_{n-1}) \cdot \bigotimes(y_1, \ldots, y_{n-1}) \geq 0$, 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 y_{H_0}, a unit vector orthogonal to H_0. For $1 \leq i \leq n$, our ith basis vector can be written as $y_i = (y_{i1}, \ldots, y_{i,n-1}, \epsilon_i)$, where by construction, each $|\epsilon_i| < 1/\|u\|$. By appropriate choice of N we can arrange that $|\epsilon_i|$ is small relative to $\|y_i\|$. Let x be the unit vector in the direction of $\bigotimes(y_1, \ldots, y_{n-1})$. Then the distance of x to y_{H_0} 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 $(x, y_1, \ldots, y_{n-1})$. By definition,

$$\begin{aligned} x \cdot \bigotimes(y_1, \ldots, y_{n-1}) &= \begin{vmatrix} x_1 & x_2 & \ldots & & x_n \\ y_{11} & y_{12} & \cdots & y_{1,n-1} & \epsilon_1 \\ \vdots & & & \vdots & \\ y_{n-1,1} & y_{n-1,2} & \cdots & y_{n-1,n-1} & \epsilon_{n-1} \end{vmatrix} \\ &= x_1 \det(M_1) - \ldots + (-1)^{n+1} x_n \det(M_n) \end{aligned} \tag{24}$$

where M_i is the $(1, i)$ minor of the matrix (expanding along the first row). Thus, $|V| = \|y_1\| \ldots \|y_{n-1}\| \cdot \|x\| = \|y_1\| \ldots \|y_{n-1}\|$. 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 $y_i^* = (y_{i1}, \ldots, y_{i,n-1})$. Since $\|x\| = 1$, $\det(M_i^*)$ is bounded by the volume of the parallelepiped with sides $y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_{n-1}$. Let $V^* = \mathrm{vol}\mathcal{P}(y_1^*, \ldots, y_{n-1}^*)$, let $\epsilon = \max_{i=1}^{n-1} \epsilon_i$ and $\alpha = \min_{i=1}^{n-1} \|y_i^*\|$. Then

cd: suppressed "the $\mathcal{U}_i$"

$$\begin{aligned} |V| &\leq |x_n V^*| + \sum_{i=1}^{n-1} |\epsilon_i Y_i| \\ &\leq |x_n V^*| + n|\frac{\epsilon V^*}{\alpha}| \\ \frac{|V|}{|V^*|} &\leq |x_n| + |\frac{n\epsilon}{\alpha}|. \end{aligned} \tag{25}$$

But $1 \leq \frac{|V|}{|V^*|}$, so $|x_n| \geq 1 - |\frac{n\epsilon}{\alpha}|$. Since $\epsilon \leq d$ and $\alpha \geq \frac{N}{2}$, we can make x_n as close to 1 as desired by appropriate choice of N. In particular, since $y_{H_0} = (0, 0, \ldots, 0, 1)$, the distance from x to the unit vector orthogonal to H_0 can be made as small as desired.

Lemma 9.6 *Assume that $b_1, ..., b_n$ is a basis of a lattice $L \subseteq \mathbb{R}^n$, $b'_1, ..., b'_n$ is its dual basis, $\|b'_i\| \leq M$ for $i = 1, ..., n$, $v \in L$, $u = \sum_{i=1}^n \beta_i b_i \in \mathbb{R}^n$, $\|u - v\|_2 < \frac{1}{2M}$ 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 $|\beta_i - \gamma_i| < \frac{1}{2}$. $|\beta_i - \gamma_i| = |b'_i \cdot v - b'_i \cdot u| = |b'_i(v - u)| < M\frac{1}{2M} = \frac{1}{2}$. ■

We want to apply Lemma 9.6 to Λ. Note that although Λ^* has a basis of length at most $2Cn^{1.5}$ we have only an LLL-reduced basis for Λ^*, of length at most $M = 2^n \mathrm{bl}(\Lambda^*)$. We compute the dual basis for this particular basis of Λ^* to get the dual basis for Λ. Suppose we have x, where $u_{H_0} \in L^*$ and x is close to u_{H_0}. The lemma says that if $\|x - u_{H_0}\| < \frac{1}{2M}$ then if we write x as a linear combination of the basis vectors for Λ and round the coefficients of these basis vectors to the nearest integers, we will obtain u_{H_0}. In order to ensure that $\|x - u_{H_0}\| \leq \frac{1}{2M}$ it is necessary and sufficient that $\sqrt{1 - x_n^2} < \frac{1}{2M}$. (The numerator $2dn$ comes from: $\alpha < N/2$ and $\epsilon < d$.) Since $|x_n| \geq 1 - |\frac{n\epsilon}{\alpha}| \geq 1 - |\frac{2dn}{N}|$, choosing $N > 16dM^2n$ suffices.

The Choice of $\mathcal{K}(n)$. The choice of $\mathcal{K}(n)$ is constrained by the requirements for Lemma 5.1 when M, the upper bound on the length of the lattice used for sampling Λ^*, is $2^n(Cn^{1.5}/\|u\|)$ (see Equation 21; the factor 2^n comes from the LLL reduction). Since the lemma is applied with offsets αv, where $\alpha \in [0, 1]$ and $\|v\| \leq N$ (N is the bound on the lengths of the vectors needed to yield a good approximation to H_0), we need that $N \leq \mathcal{K}(n)^{1/4}$.

10 Tightening the Approximation Factor

In this section we improve the approximation factor to $\tilde{O}(n^2)$. Along the way, we will reduce the size of the public key from $\tilde{O}(n^4)$ to $O(n^4)$.

We use several tricks. However, we will not change the way in which the vertices $V = (v_0, \ldots, v_\ell)$ are chosen. As above, we only consider the case in which each v_i has length at most $M \leq \log^2 n\sqrt{n + \ell}\mathcal{K}(n)$.

1. Make $\mathcal{P}$ very close to orthogonal. By doing this we ensure that the width of $\mathcal{P}$ is essentially the length of the shortest p_i. This eliminates a factor of n from the coefficients λ_i when the ciphertext, before modding out by $\mathcal{P}$, is $\sum_i \lambda_i p_i + \mathbf{o}$.

2. Make $\mathcal{P}$ very large. By doing this we can make the coefficients of $\sum_i b_i v_i$ constant. In order to keep m' small we will need to increase the lengths of the d_i accordingly. The coefficients of $\sum_j 2\delta_j d_j$ will be polylogarithmic in n.

3. Modify the encryption procedure so that to encrypt a message $b_0 b_1, \ldots b_\ell$ one first chooses an $(\ell + 1)$-bit random pad x and then sends the encryption of x together with $x \oplus b_0 b_1 \ldots b_\ell$. This reduces the distance of the encrypted pad, before modding out by $\mathcal{P}$, from the worst-case bound $n\rho$ to an expected $\sqrt{n}\rho$ (and an overwhelmingly likely $\tilde{O}(\sqrt{n}\rho)$).

4. Before modding out by $\mathcal{P}$ we have a ciphertext
$$\sum_{i=0}^{\ell} x_i v_i + \sum_{j=1}^{m'} \delta_j d_j = \sum_{i=1}^{n+\ell} \lambda_i p_i + \mathbf{o}$$
Refine the analysis of the harm caused by the coefficients λ_i to reflect the fact that $\mathrm{dist}(\sum_i p_i, \mathcal{H}_u)$ is very likely to be close to $\sqrt{n}\rho$ rather than the worst-case $n\rho$. This will translate to an expected $O(\rho\sqrt{2nm'})$ distance of the ciphertext to each $\mathcal{H}_u$, after modding out by $\mathcal{P}$.

We now discuss these in more detail.

Making $\mathcal{P}$ Nearly Orthogonal and Very Large. Note that if we only make $\mathcal{P}$ close to orthogonal, and follow none of the additional suggestions, we reduce the approximation factor by n. There are several ways in which this may be done. We will describe one that only doubles the size of the public key and the ciphertext. In passing, then, this gives us a cryptosystem with a linear amortized plaintext to ciphertext blowup, public keys of size $\tilde{O}(n^4)$, and an approximation factor of $\tilde{O}(n^{3.5})$.

In general, to obtain a nearly-orthogonal parallelepiped, we choose each p_i to be close to $R\mathbf{e_i}$, for a suitable choice of R and for $i = 1, \ldots, n+\ell$. Since during the reduction we need to be able to find lattice points in Λ^* close to $R\mathbf{e_i}$, our notion of "close" must be at least on the order of 2^n, which is roughly $\mathcal{K}(n)^{1/4}$. We need this quantity to be small compared to R. We can achieve this by taking R to be $\mathcal{K}(n)^2$. Each p_i will be obtained by taking a sample of $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_\ell}^{n+\ell,\mathcal{K}(n)} + \mathcal{N}_{n+\ell}(0,\rho^2)$ and adding it to $R\mathbf{e_i}$.

The d_j, $1 \le j \le m'$, will be chosen from $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_\ell}^{n+\ell,R/\sqrt{n+\ell}} + \mathcal{N}_{n+\ell}(0,\rho^2)$, so as to have expected length approximately the same as the lengths of the sides of the parallelepiped. This allows us to apply the argument in Section 7, ensuring that $m' \in \tilde{O}(m')$. This already ensures that $\sum_{i=0}^{\ell} b_i v_i = \sum_{i=1}^{n+\ell} \lambda_i p_i + \mathbf{o}$ will have coefficients λ_i bounded in magnitude by 1, since the sides of the parallelepiped have length much larger than n times the maximum length v_i.

Analysis of $\mathrm{dist}(\sum_{i=0}^{\ell} x_i v_i + 2\sum_{j=0}^{m'} \delta_j d_j, \mathcal{H}_u)$**.** We first give a simple argument that with high probability over V and D, once we fix V, D, random subset sums are unlikely to be far from $\mathcal{H}_u$.

Let X_i, $i = 1, \ldots, N$, be i.i.d. random variables with distribution $\mathcal{N}_1(0,1)$. Let x_i, $i = 1, \ldots, N$, be i.i.d. random variables with uniform distribution on $\{0,1\}$. We say that $x = (x_1, \ldots, x_N)$ is *bad for* $X = (X_1, \ldots, X_N)$, denoted $\mathrm{Bad}(x, X)$, if $|\sum_{i=1}^{N} x_i X_i| > \eta\sqrt{N}$ for a quantity η to be determined later. Let $I(x, X)$ be the indicator random variable with value 1 if $\mathrm{Bad}(x, X)$ and 0 otherwise. Let
$$f(X) = \mathrm{Pr}_x[\mathrm{Bad}(x, X)] = \frac{1}{2^N} \sum_{x \in \{0,1\}^N} I(x, X)\,.$$

For $y = (y_1, \ldots, y_N) \in \mathrm{I\!R}^N$, let $\varphi(y)$ denote the probability density function $\prod_{i=1}^{N} g_1^1(y_i)$. Then
$$E_X[f(X)] \quad = \quad \int_{X \in \mathrm{I\!R}^N} \mathrm{Pr}_x[\mathrm{Bad}(x, X)]\, \varphi(X) dX$$

$$
\begin{aligned}
&= \int_{X\in \mathbb{R}^N} \frac{1}{2^N} \sum_{x\in\{0,1\}^N} I(x,X)\varphi(X)dX \\
&= \frac{1}{2^N} \sum_{x\in\{0,1\}^N} \int_{X\in \mathbb{R}^N} I(x,X)\varphi(X)dX
\end{aligned}
$$

Let $H(x)$ denote the Hamming weight of x. For a fixed sequence x the integral

$$\int_{X\in \mathbb{R}^N} I(x,X)\varphi(X)dX$$

is just the probability of the event that the absolute value of the sum of $H(x) \leq N$ independent standard normals exceeds $\eta\sqrt{N}$. When η is sufficiently large, eg, $\eta \geq \log^2 N$, this probability is negligible. Summing over all possible sequences x we get:

$$
\begin{aligned}
E_X[f(X)] &= \frac{1}{2^N} \sum_{x\in\{0,1\}^N} \nu(N) \\
&= \nu(N)
\end{aligned}
$$

We may now apply the Markov inequality:

$$\mathrm{Pr}_X[f(X) > z] < E_X[f(X)]/z$$

to conclude that with all but negligible probability over choice of X, $f(X)$ is negligible. For example, we may take

$$
\begin{aligned}
z &= \sqrt{E_X[f(X)]} \\
&\Rightarrow \ E_X[f(X)]/z = \sqrt{E_X[f(X)]}
\end{aligned}
$$

This is $\sqrt{\nu(N)}$, a quantity still negligible in N.

Since for $i = 0, \ldots, \ell$ the signed distance of v_i to $\mathcal{H}_u$ has distribution $\mathcal{N}_1(0,\rho^2)$, and since the same is true for the signed distance of d_j to $\mathcal{H}_u$, $j = 1, \ldots, m'$, the argument implies that with all but negligible probability over V, D, the unsigned distance $\mathrm{dist}(\sum_{i=0}^{\ell} x_i v_i + 2\sum_{j=0}^{m'} \delta_j d_j, \mathcal{H}_u)$ is bounded by $\log^2(n+m')\sqrt{n+m'}\,\rho$ with all but negligible probability over the choice of $x_0, \ldots, x_\ell, \delta_1, \ldots, \delta_{m'}$.

Next, we modify the distribution from which we choose the d_j. The point of the modification is to reduce the size of m' to $O(n^4)$. The modification will cause the coefficients λ_i to increase slightly. These two effects will more or less cancel out in the approximation factor, while still permitting the slightly smaller public key. The change will not affect the analysis of $\mathrm{dist}(\sum_{i=0}^{\ell} x_i v_i + 2\sum_{j=0}^{m'} \delta_j d_j, \mathcal{H}_u)$.

Intuitively, in Step 0 of the chain we want to choose $d_1, \ldots, d_{m'}$ according to a Gaussian with parameter exceeding the smoothing parameter of the lattice $L(p_1, \ldots, p_{n+\ell})$. Since we may assume that the p_i, $1 \leq i \leq n+\ell$, are all of length at most $2R$, in the public key of the cryptosystem, the d_j will be drawn from $\mathcal{H}\mathrm{Samp}_{u_0,\ldots,u_\ell}^{n+\ell,\gamma(n)2R} + \mathcal{N}_{n+\ell}(0,\rho^2)$, where $\gamma(n) \in \sqrt{\omega(\log n)}$. At Step 0 of the chain the d_j, $1 \leq j \leq m'$, will therefore have distribution negligibly close to uniform modulo $\mathcal{P}$, and we may apply Lemma 7.1 to obtain an upper bound of $m' \in O(n^2)$.

Refined Analysis of the Damage Caused by the λ_i. As noted above, we may assume

$$\mathrm{dist}(\sum_{i=0}^{\ell} b_i v_i + \sum_{j=1}^{m'} \delta_j d_j, \mathcal{H}_u) < \log^2(n+m')\sqrt{n+m'}\,\rho \in \tilde{O}(n\rho). \tag{26}$$

On the other hand, there is the concern that the distance from $\mathcal{H}_u$ caused by the coefficients of the p_i – of which there are $n+\ell$ – in

$$\sum_{i=0}^{\ell} b_i v_i + \sum_{j=1}^{m'} \delta_j d_j = \sum_{i=1}^{n+\ell} \lambda_i p_i + \mathbf{o} \tag{27}$$

could be as bad as $(n+\ell)\rho$ times the maximum coefficient. However, this is very unlikely, for the following reason. For $i = 1, \ldots, n+\ell$, let ψ_i denote the *signed* distance of p_i from $\mathcal{H}_u$.

Then the distance introduced by modding out by $\mathcal{P}$ is $\sum_{i=1}^{n+\ell} \lambda_i \psi_i$. The ψ_i are not completely independent of the λ_i since if we tile the space by copies of the parallelepiped $\mathcal{P}$ then, if a point before perturbation is near to the boundary of a cell of the tiling, the perturbation may take it to another cell of the tiling and so change λ_i. However the probability (for the selection of the point) that this happens is so small that we may assume this does not occur, and the consequences will remain valid with high probability.

Writing $d_j = \sum_{i=1}^{n+\ell} \alpha_i p_i + \mathbf{o}$, for integers $\alpha_1, \ldots, \alpha_{n+\ell}$, since the p_i are almost mutually orthogonal the α_i are nearly independent (this can be made rigorous by showing that the distribution on $(\alpha_1, \ldots, \alpha_{n+\ell})$ is extremely close to the distribution on coefficients of $(R\mathbf{e_1}, \ldots, R\mathbf{e_{n+\ell}})$ if we express d_j as a linear combination of these vectors) and each has a normal distribution $\mathcal{N}_1(0, \gamma(n))$. Since the d_j are mutually independent, their sums have coefficients distributed as $\mathcal{N}_1(0, \tilde{O}(m'))$.

In addition the variance of the errors ψ_i is ρ^2, and indeed the ψ_i are distributed according to $\mathcal{N}_1(0, \rho^2)$, so the variance of the products $\lambda_i\psi_i$ is $\tilde{O}(m'\rho^2)$ and, for all i, i', $\lambda_i\psi_i$ is independent of $\lambda_{i'}\psi_{i'}$. We therefore have a sum of $n+\ell$ independent variables of variance $\tilde{O}(m'\rho^2)$ and mean 0, yielding a total variance of $(n+\ell)\tilde{O}(m'\rho^2) \leq \tilde{O}(2nm'\rho^2)$, and a standard deviation of $\tilde{O}(\sqrt{2nm'}\rho) = \tilde{O}(n^{1.5}\rho)$.

Putting the Pieces Together. The constraint on ρ necessary for correct decoding is that it will be at least twice the error introduced by the subset sums and modding by $\mathcal{P}$. It is clear that for some constants c, c', taking $\rho = c'/n^{1.5}\log^c n$ satisfies all these conditions, and this is precisely what is needed for semantic security based on the assumed worst-case hardness of the $\tilde{O}(n^2)$-unique shortest vector problem for lattices.

11 Conclusions

In this work we have extended and generalized our original public-key cryptosystem, the first with a proved worst-case/average-case equivalence. We have made two contributions: an algorithmic innovation, generalizing the original construction so as to permit good amortized plaintext to ciphertext expansion, and a method of selecting the public keys that simplifies the proof of correctness of the generalization. We have also exploited recent work on lattices to reduce some of the parameters.

In terms of the four parameters: approximation factor, size of public key, size of ciphertext, and plaintext to ciphertext expansion, our cryptosystem is comparable to that of Regev [14], with a slightly worse approximation factor ($\tilde{O}(n^2)$ vs. $n^{1.5}$) but a much better amortized expansion.

Better bounds are achieved by Regev in [15]. In particular, he obtains a public key of size $\tilde{O}(n^2)$ instead of $O(n^4)$ in our construction, and he achieves an almost comparable expansion factor, without amortization. The cryptosystem in [15] relies on an assumption regarding the quantum complexity of the unique shortest vector problem. If the quantum complexity of this problem is low, then there is room for our construction – at least until quantum computers are actually manufactured – and this is where our non-quantum proof of security is relevant.

The large size of the public key comes from two not unrelated sources: m', the number of "dust" points, and the value of $\mathcal{K}(n)$. If in the future a method is developed for finding a basis of length subexponential in $\mathrm{bl}(\Lambda^*)$, but the $\tilde{O}(n^2)$-unique shortest vector problem for lattices is still hard, then we can reduce $\mathcal{K}(n)$ and, in consequnce, m' as well (recall that the requirement is (very) roughly that $2^{m'} > (\mathcal{K}(n)2^{\mathbf{p}})^n$, or $m' > n(\log \mathcal{K}(n) + \mathbf{p})$). For example, if finding a basis of length $\exp(\log^c n)\mathrm{bl}(\Lambda^*)$ is feasible for some $c > 2$, but finding an $n^{1.5}$-unique shortest vector is not, then by choosing $\mathbf{p} \in \mathrm{polylog}(n)$ we can get by with $m' \in \tilde{O}(n)$. This would shrink the public key and the ciphertext by a factor of almost n^2, and also permit the tighter approximation factor of $\tilde{O}(n^{1.5})$.

References

[1] M. Ajtai, Generating Hard Instances of Lattice Problems. *Quaderni di Matematica, Publ. Seconda Universita di Napoli, Vol. 13, "Complexity of Computations and Proofs"*, (2005), pp 1-32 A preliminary version appeared in Proceedings 28th Annual ACM Symposium on Theory of Computing, 1996.

[2] M. Ajtai and C. Dwork, A Public-Key Cryptosystem with Worst-Case/Average-Case Equivalence, *Proc. STOC'97*

[3] R. Ash, *Basic Probability Theory*, Wiley, 1970

[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] J-Y. Cai, A Relation of Primal-Dual Lattices and the Complexity of Shortest Lattice Vector Problem, *Theoretical Computer Science 207*(1), 1998.

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

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

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

[9] S. Goldwasser and S. Micali, Probabilistic Encryption. *J. Comput. Syst. Sci. 28*(2), pp. 270–299, 1984.

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

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

[12] D. Micciancio, SiComp 2004

[13] D. Micciancio and O. Regev Worst-case to Average-case Reductions based on Gaussian Measures, Proc. of FOCS 2004. To appear, *SiComp.*

[14] O. Regev, New Lattice Based Cryptographic Constructions Oded Regev Journal of the ACM 51(6), pp. 899-942, 2004. Preliminary version in Proc. of STOC 2003.

[15] O. Regev, On Lattices, Learning with Errors, Random Linear Codes, and Cryptography, Proc. of STOC 2005.

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

ECCC ISSN 1433-8092

http://eccc.hpi-web.de/