A Lattice-Based Public-Key Cryptosystem

Jin-Yi Cai* and Thomas W. Cusick

[1] Department of Computer Science
State University of New York at Buffalo, Buffalo, NY 14260
cai@cs.buffalo.edu
[2] Department of Mathematics
State University of New York at Buffalo, Buffalo, NY 14260
cusick@acsu.buffalo.edu

Abstract. Ajtai recently found a random class of lattices of integer points for which he could prove the following worst-case/average-case equivalence result: If there is a probabilistic polynomial time algorithm which finds a short vector in a random lattice from the class, then there is also a probabilistic polynomial time algorithm which solves several problems related to the shortest lattice vector problem (SVP) in any n-dimensional lattice. Ajtai and Dwork then designed a public-key cryptosystem which is provably secure unless the worst case of a version of the SVP can be solved in probabilistic polynomial time. However, their cryptosystem suffers from massive data expansion because it encrypts data bit-by-bit. Here we present a public-key cryptosystem based on similar ideas, but with much less data expansion.
Keywords: Public-key cryptosystem, lattice, cryptographic security.

1 Introduction

Since the origin of the idea of public-key cryptography, there have been many public-key techniques described in the literature. The security of essentially all of these depends on certain widely believed but unproven mathematical hypotheses. For example, the well-known RSA public-key cryptosystem relies on the hypothesis that it is difficult to factor a large integer n which is known to be a product of two large primes. This hypothesis has been extensively studied, but there is still no proof that for a typical such n, the prime factors cannot be found in less than k steps, where k is a very large number. From a computational complexity point of view, we generate a specific instance of a problem in NP (together with a solution, which is kept secret) and we rely on the belief that the problem is difficult to solve.

Apart from the lack of proof that any of these problems is really hard, i.e., there exists no efficient algorithm that will solve the problem in all cases, there is another serious issue. The mathematical hypothesis that these problems are difficult to solve really means difficult to solve in *the worst case*, but the security

* Research supported in part by NSF grant CCR-9634665 and an Alfred P. Sloan Fellowship.

S. Tavares and H. Meijer (Eds.): SAC'98, LNCS 1556, pp. 219–233, 1999.

of the cryptographic algorithms depends more on the difficulty of *the average case*. For example, even if one day factoring is proved to be unsolvable in probabilistic polynomial time, to the users of the RSA system, there is no guarantee that the key they are actually using is hard to factor. To use these protocols, one must be able to generate specific instances of the problem which should be hard to solve. But typically there is no way to just generate known hard instances. One way to do this is to generate random instances of the problem, and hope that such instances are as hard on the average as in the worst case. However this property is known to be not true for a number of NP-hard problems.

Recently Ajtai [1] proved that certain lattice problems related to SVP have essentially the same average case and worst case complexity, and both are conjectured to be extremely hard. This development raises the possibility of public-key cryptosystems which will have a new level of security. Already Ajtai and Dwork [3] have proposed a public-key cryptosystem which has a provable worst-case/average-case equivalence. Specifically, the Ajtai-Dwork cryptosystem is secure unless the worst case of a certain lattice problem can be solved in probabilistic polynomial time.

Goldreich, Goldwasser and Halevi [11] have also given a public-key cryptosystem which depends on similar lattice problems related to SVP as in [1]. Unlike the work of [3], however, their method uses a trapdoor one-way function and also lacks a proof of worst-case/average-case equivalence.

The cryptosystems of [3] are unfortunately far from being practical. All of them encrypt messages bit-by-bit and involve massive data expansion: the encryption will be at least a hundred times as long as the message. (Note: In a private communication, Ajtai has informed us that this data expansion problem is being addressed by the authors of [3] as well.) In this paper we propose a public-key cryptosystem, based on the ideas of [1] and [3], which has much less data expansion. Messages are encrypted in blocks instead of bit-by-bit. We offer some statistical analysis of our cryptosystem. We also analyze several attacks on the system and show that the system is secure against these attacks. Whether there is a provable worst-case/average-case equivalence for this system is open.

2 Lattice problems with worst-case/average-case equivalence

Here we briefly define the terms for lattice problems, and describe the results of Ajtai [1] and some improvements.

Notation. $\mathbf{R}$ is the field of real numbers, $\mathbf{Z}$ is the ring of integers, $\mathbf{R}^n$ is the space of n-dimensional real vectors $a = \langle a_1, \ldots, a_n \rangle$ with the usual dot product $a \cdot b$ and Euclidean norm or length $\|a\| = (a \cdot a)^{1/2}$. $\mathbf{Z}^n$ is the set of vectors in $\mathbf{R}^n$ with integer coordinates, $\mathbf{Z}^+$ is the positive integers and $\mathbf{Z}_q$ is the ring of integers modulo q.

Definition. If $A = \{a_1, \ldots, a_n\}$ is a set of linearly independent vectors in $\mathbf{R}^n$, then we say that the set of vectors

$$\{\textstyle\sum_{i=1}^{n} k_i a_i : \ k_1, \ldots, k_n \in \mathbf{Z}\}$$

is a lattice in $\mathbf{R}^n$. We will denote the lattice by $L(A)$ or $L(a_1, \ldots, a_n)$. We call A a basis of the lattice. We say that a set in $\mathbf{R}^n$ is an n-dimensional lattice if there is a basis V of n linearly independent vectors such that $L = L(V)$. If $A = \{a_1, \ldots, a_n\}$ is a set of vectors in a lattice L, then we define the length of the set A by $\max_{i=1}^{n} \|a_i\|$. $\lambda_1(L) = \min_{0 \neq v \in L} ||v||$.

A fundamental theorem of Minkowski is the following:

Theorem 1 (Minkowski). *There is a universal constant γ, such that for any lattice L of dimension n, $\exists v \in L$, $v \neq 0$, such that*

$$||v|| \leq \gamma \sqrt{n} \det(L)^{1/n}.$$

The determinant $\det(L)$ of a lattice is the volume of the n-dimensional fundamental parallelepiped, and the absolute constant γ is known as Hermite's constant.

Minkowski's theorem is a pure existence type theorem; it offers no clue as to how to find a short or shortest non-zero vector in a high dimensional lattice. To find the shortest non-zero vector in an n-dimensional lattice, given in terms of a basis, is known as the Shortest Vector Problem (SVP). There are no known efficient algorithms for finding the shortest non-zero vector in the lattice. Nor are there efficient algorithms to find an approximate short non-zero vector, or just to approximate its length, within any fixed polynomial factor in its dimension n. This is still true even if the shortest non-zero vector v is unique in the sense that any other vector in the lattice whose length is at most $n^c\|v\|$ is parallel to v, where c is an absolute constant. In this case we say that v is unique up to a polynomial factor.

The best algorithm to date for finding a short vector in an arbitrary lattice in $\mathbf{R}^n$ is the L^3 algorithm of A.K. Lenstra, H.W. Lenstra and L. Lovász [14]. This algorithm finds in deterministic polynomial time a vector which differs from the shortest one by at most a factor $2^{(n-1)/2}$. C.P. Schnorr [16] proved that the factor can be replaced by $(1+\epsilon)^n$ for any fixed $\epsilon > 0$. However Schnorr's algorithm has a running time with $1/\epsilon$ in the exponent.

Regarding computational complexity, Ajtai [2] proved that it is NP-hard to find the shortest lattice vector in Euclidean norm, as well as approximating the shortest vector length up to a factor of $1 + \frac{1}{2^{n^k}}$. In a forthcoming paper [6], Cai and Nerurkar improve the NP-hardness result of Ajtai [2] to show that the problem of approximating the shortest vector length up to a factor of $1 + \frac{1}{n^\varepsilon}$, for any $\varepsilon > 0$, is also NP-hard. This improvement also works for all l_p-norms, for $1 \leq p < \infty$. Prior to that, it was known that the shortest lattice vector problem is NP-hard for the l_∞-norm, and the nearest lattice vector problem is NP-hard under all l_p-norms, $p \geq 1$ [12,17]. Even finding an approximate solution to within any constant factor for the nearest vector problem for any l_p-norm is NP-hard [4]. On the other hand, Lagarias, Lenstra and Schnorr [13] showed that the approximation problem (in l_2-norm) within a factor of $O(n)$ *cannot* be NP-hard, unless NP = coNP. Goldreich and Goldwasser showed that approximating the shortest lattice vector within a factor of $O(\sqrt{n/\log n})$ is not NP-hard assuming the polynomial time hierarchy does not collapse [9]. Cai

showed that finding an $n^{1/4}$-unique shortest lattice vector is not NP-hard unless the polynomial time hierarchy collapses [7].

What is most striking is a recent result of Ajtai [1] establishing the first explicit connection between the worst-case and the average-case complexity of the problem of finding the shortest lattice vector or approximating its length. The connection factor in the Ajtai connection has been improved in [5]. Ajtai defined a class of lattices in $\mathbf{Z}^m$ so that if there is a probabilistic polynomial time algorithm which finds a short vector in a random lattice from the class with a probability of at least $1/n^{O(1)}$, then there is also a probabilistic polynomial time algorithm which solves the following three lattice problems in *every* lattice in $\mathbf{Z}^n$ with a probability exponentially close to 1:

(P1) Find the length of a shortest non-zero vector in an n-dimensional lattice, up to a polynomial factor.
(P2) Find the shortest non-zero vector in an n-dimensional lattice where the shortest vector is unique up to a polynomial factor.
(P3) Find a basis in an n-dimensional lattice whose length is the smallest possible, up to a polynomial factor.

The lattices in the random class are defined modulo q (q is an integer depending only on n, as described below), that is, if two integer vectors are congruent modulo q then either both of them or neither of them belong to the lattice. More precisely, if $\nu = \{u_1, \ldots, u_m\}$ is a given set of vectors in $\mathbf{Z}_q^n$ then the lattice $\Lambda(\nu)$ is the set of all integer vectors $\langle h_1, \ldots, h_m \rangle$ so that

$$\sum_{i=1}^{m} h_i u_i \equiv 0 \pmod{q}.$$

For a fixed n, m and q, the probability distribution over the random class is defined by uniformly choosing a sequence of integer vectors $\langle u_1, \ldots, u_m \rangle$.

For a given n, the parameters m and q are defined by $m = [c_1 n]$ and $q = [n^{c_2}]$, where c_1 and c_2 are suitable constants.

The problem of finding a short vector in a lattice from the random class is a Diophantine problem. Questions of this type date back to Dirichlet's 1842 theorem on simultaneous Diophantine approximation. From this point of view the problem can be stated in the following way, which does not involve any explicit mention of lattices, as pointed out in [10].

(A1) Given $n, m = [c_1 n], q = [n^{c_2}]$ and an n by m matrix M with entries in $\mathbf{Z}_q$, find a non-zero vector x so that $Mx \equiv 0 \pmod{q}$ and $\|x\| < n$.

Minkowski's theorem guarantees the existence of such short vectors x. Of course if the condition on $\|x\|$ is removed, then the linear system $Mx \equiv 0 \pmod{q}$ can be solved in polynomial time.

The theorem in [1] reduces the worst-case complexity of each of the problems **(P1)**, **(P2)**, **(P3)** to the average case complexity of **(A1)**. Currently the best bounds that can be achieved are stated below [5], [8]:

Theorem 2. *[5] For any constant $\epsilon > 0$, if there exists a probabilistic polynomial time algorithm $\mathcal{A}$ such that, for a given random lattice $\Lambda(\nu)$, where*

$\nu = (u_1, \ldots, u_m) \in \mathbf{Z}_q^{n \times m}$ is uniformly chosen, $q = \Theta(n^3)$ and $m = \Theta(n)$, $\mathcal{A}$ will find a vector of the lattice $\Lambda(\nu)$ of length $\leq n$ with probability $\frac{1}{n^{O(1)}}$, then, there also exists a probabilistic polynomial time algorithm $\mathcal{B}$ which for any given lattice $L = L(a_1, \ldots, a_n)$ by a basis $a_1, \ldots, a_n \in \mathbf{Z}^n$, outputs another basis for L, $b_1, \ldots, b_n$, so that,

$$\max_{i=1}^{n} \|b_i\| \leq \Theta(n^{3.5+\epsilon}) \min_{\text{all bases } b'_1, \ldots, b'_n \text{ for } L} \max_{i=1}^{n} \|b'_i\|.$$

Theorem 3. *[8] Under the same hypothesis, there exists a probabilistic polynomial time algorithm $\mathcal{C}$ which for any given lattice $L = L(a_1, \ldots, a_n)$ by a basis will*

- *compute an estimate of $\lambda_1 = \lambda_1(L)$ up to a factor $n^{4+\epsilon}$, i.e., compute a numerical estimate $\tilde{\lambda}_1$, such that*

$$\frac{\lambda_1}{n^{4+\epsilon}} \leq \tilde{\lambda}_1 \leq \lambda_1;$$

- *find the unique shortest vector if it is an $n^{4+\epsilon}$-unique shortest vector.*

3 A new cryptosystem

Here we present the design of a new cryptosystem, which is based on the difficulty of finding or approximating SVP, even though no specific lattices are defined. The secret key in the new system is a vector u chosen with uniform distribution from the unit sphere $S^{n-1} = \{x \mid ||x|| = 1\}$, and a random permutation σ on $m+1$ letters. By allowing an exponentially small round-off error, we may assume that the coordinates of u are rational numbers whose denominators are bounded by some very large integer, exponential in n. Let $m = [cn]$ for a suitable absolute constant $c < 1$. For definiteness set $c = 1/2$. Let $H_i = \{v : \ v \cdot u = i\}$ denote the hyperplanes perpendicular to u. The public key in the new system is a parameter $b > 0$ and a set $\{v_{\sigma(0)}, \ldots, v_{\sigma(m)}\}$ of rational vectors, where each v_j is in one of the hyperplanes H_i for some $i \in \mathbf{Z}^+$, say $v_j \cdot u = N_j \in \mathbf{Z}^+$. We choose a sequence of numbers N_j so that it is superincreasing, that is

$$N_0 > b \text{ and } N_i > \textstyle\sum_{j=0}^{i-1} N_i + b \text{ for each } i = 1, 2, \ldots, m.$$

Binary plaintext is encrypted in blocks of $m+1$ bits. If $P = (\delta_0, \ldots, \delta_m)$ is a plaintext block ($\delta_i = 0$ or 1), then P is encrypted as a random perturbation of $\sum_{i=0}^{m} \delta_i v_{\sigma(i)}$. More precisely, the sender picks a uniformly chosen random vector r with $||r|| \leq b/2$. Then the ciphertext is

$$\sum_{i=0}^{m} \delta_i v_{\sigma(i)} + r.$$

Decryption is accomplished by using the secret key u to compute the following inner product

$$\begin{aligned} S &= u \cdot \left(\sum_{i=0}^{m} \delta_i v_{\sigma(i)} + r \right) \\ &= \sum_{i=0}^{m} \delta_i (u \cdot v_{\sigma(i)}) + u \cdot r \\ &= \sum_{i=0}^{m} \delta_i N_{\sigma(i)} + u \cdot r \\ &= \sum_{i=0}^{m} \delta_{\sigma^{-1}(i)} N_i + u \cdot r. \end{aligned}$$

Since the N_i are superincreasing, we can use the greedy algorithm to efficiently recover the $\delta_{\sigma^{-1}(i)}$ from S, and then use the secret σ to recover δ_i. More precisely, if $\delta_{\sigma^{-1}(m)} = 1$, then $S \geq N_m - b/2$, and if $\delta_{\sigma^{-1}(m)} = 0$, then $S \leq N_0 + N_1 + \ldots + N_{m-1} + b/2$. Since $N_m > \sum_{i=0}^{m-1} N_i + b$, with the secret key one can discover whether $\delta_{\sigma^{-1}(m)} = 0$ or 1. Substituting S by $S' = S - \delta_{\sigma^{-1}(m)} N_m$, this process can be continued until all $\delta_{\sigma^{-1}(i)}$ are recovered. Then using the secret permutation σ, one recovers $\delta_0, \delta_1, \ldots, \delta_m$.

Thus decryption using u and σ involves an easy instance of a knapsack problem. As summarized in the article of Odlyzko [15], essentially all suggestions for cryptosystems based on knapsack problems have been broken. Here, however, the easy knapsack problem appears to have no bearing on the security of the system, since it appears that one must first search for the direction u.

The new cryptosystem has similarities with the third version of the Ajtai-Dwork cryptosystem (see [3]), but in the new system $m + 1 = O(n)$ bits of plaintext are encrypted to an n-dimensional ciphertext vector, instead of just one bit of plaintext.

We have not specified the distribution of the v_i, aside from its inner product with u being superincreasing. The following distribution has a strong statistical indistinguishability from $m + 1$ independent uniform samples of the sphere. Let M be a large integer, say, $M \gg 2^n$. Choose any $b' > b$. For analysis purposes we will normalize by denoting v_i/M as v_i. For each i, $0 \leq i \leq m$, let $v_i = \frac{2^i b'}{M} u + \sqrt{1 - \frac{2^{2i} b'^2}{M^2}} \rho_i$, where the ρ_i's are independently and uniformly distributed on the $(n-2)$-dimensional unit sphere orthogonal to u. Note that each $||v_i|| = 1$, after normalization. We denote this distribution by D. We note that

$$u \cdot v_i - u \cdot \left(\sum_{j=0}^{i-1} v_j \right) = \frac{2^i b'}{M} > \sum_{j=0}^{i-1} \frac{2^j b'}{M} = \frac{b'}{M} > \frac{b}{M}.$$

How secure is this new cryptosystem? We do not have a proof of worst-case/average-case equivalence. We can discuss several ideas for attacks that do

not seem to work. The following discussion will also explain some of the choices made in the design of the cryptosystem.

We will first show that if we did not employ the random permutation σ, rather we publish as public key the unpermuted vectors $v_0, \ldots, v_m$, then there is an attack based on linear programming that will break the system in polynomial time.

The attack works as follows: From the given vectors $v_0, \ldots, v_m$ we are assured that the following set of inequalities defines a non-empty convex body containing the secret vector u.

$$\begin{aligned} v_0 \cdot x &> b \\ v_1 \cdot x &> v_0 \cdot x + b \\ v_2 \cdot x &> (v_0 + v_1) \cdot x + b \\ &\vdots \\ v_m \cdot x &> (v_0 + v_1 + \cdots + v_{m-1}) \cdot x + b \end{aligned}$$

Using linear programming to find a *feasible* solution to this convex body, we can compute in polynomial time a vector $\tilde{u}$ satisfying all the inequalities. Even though $\tilde{u}$ may not be equal to u, as along as $\tilde{u}$ satisfies the above set of inequalities, it is as good as u itself to decrypt the message $\sum_{i=0}^{m} \delta_i v_i + r$. Hence, the permutation σ is essential to the security of the protocol.

Next, let's consider the addition of the random perturbation r. This is to guard against an attack based on linear algebra, which works as follows.

Assume the message $w = \sum_{i=0}^{m} \delta_i v_{\sigma(i)}$ were sent without the perturbation vector r. Then this vector is in the linear span of $\{v_{\sigma(0)}, v_{\sigma(1)}, \ldots, v_{\sigma(m)}\}$, which is most likely to be linearly independent, by the way these v_i's are chosen. Then one can solve for the $m + 1 < n$ coefficients x_i in $w = \sum_{i=0}^{m} x_i v_{\sigma(i)}$. These coefficients are unique by linear independence, thus $x_i = \delta_i$, and we recover the plaintext.

The addition of the random perturbation r renders this attack ineffective, since with probability very near one, $w = \sum_{i=0}^{m} \delta_i v_{\sigma(i)} + r$ is not in the linear span of $\{v_{\sigma(0)}, v_{\sigma(1)}, \ldots, v_{\sigma(m)}\}$, which is of dimension at most $m + 1$. (If r were truly uniformly random from the ball $||x|| \leq b/2$, then the probability that w belongs to the lower dimensional linear span is zero; if r is chosen with rational coordinates with exponentially large denominator, then this probability is exponentially small.) When the vector w is not in the linear span, to recover the coefficients δ_i appears to be no easier than the well known nearest lattice vector problem, which is believed to be intractable.

Finally if the lengths of the vectors v_i are not kept essentially the same, there can be statistical leakage of information (see Section 4). However, suppose the v_i are all roughly the same length, then the number of message bits $m = cn$ should be less than n. If m were equal to cn, for a constant $c > 1$, then there is the following cryptanalytic attack.

Suppose $\|v_i\| \approx V$ for each i and define numbers Q_i by

$$Q_i = \frac{|u \cdot v_i|}{\|u\|\|v_i\|} \approx \frac{N_i}{\|u\|V}.$$

Since the integers N_i are superincreasing, we can show that for all $i < m - 3\log n$, $Q_i < Q_m/n^3$. In fact, let $m' = m - 3\log n$, then we can inductively prove that $Q_{m'+j} > 2^j Q_{m'-1} \geq 2^j Q_i$ for all $i < m'$. Thus for each $i < m'$ we have $Q_i < Q_m/n^3$. We will say that these Q_i are "unusually small" (compared to $\max Q_j$). Of course one cannot compute the Q_i's since one is given only the permuted ordering by σ and u is secret.

The attack begins with the selection of a random subset of $n-1$ vectors v_i. If we get all $n-1$ vectors having an unusually small dot product with the secret vector u, then the normal vector perpendicular to all these $n-1$ vectors will be a good approximation to u. From this one can break the system. We show next that with non-trivial probability $1/n^{O(1)}$, all $n-1$ vectors have an unusually small dot product with the secret vector u. This is at least

$$\frac{\binom{cn-3\log n}{n-1}}{\binom{cn}{n-1}} \geq \left(1 - \frac{3\log n}{cn-n}\right)^n \approx n^{-\frac{3}{c-1}}.$$

Thus one can try for a polynomial number of times, and with high probability one will find such a set of $n-1$ vectors and break the system. This attack will not work if $m = n/2$.

4 Statistical analysis

It is clear from the discussion that the secret permutation σ as well as the random perturbation r are both necessary. With a secret permutation σ, however, an adversary may still attempt to find or approximate the secret vector u. In this section, the random perturbation r does not play an essential role in the analysis; it is easier to discard r in the following analysis, which is essentially the same with r, although a little less clean. Thus we will carry out the following analysis with $b = 0$ and $r = 0$.

A natural attack is to gather statistical information by computing some values associated with the vectors $v_{\sigma(0)}, \ldots, v_{\sigma(m)}$ which are invariant under the permutations. It is conceivable, for example, that $\sum_{i=0}^{m} v_i = \sum_{i=0}^{m} v_{\sigma(i)}$ might have a non-trivial correlation with the secret direction u since each v_i has a positive component in the u direction. We will show that if we did not choose our distribution for the v_i's carefully, then indeed this attack may succeed; but the distribution D appears to be secure against this attack.

Consider again the structural requirement that $v_0 \cdot u > 0$, and $v_i \cdot u > (v_0 + \cdots + v_{i-1}) \cdot u$. A natural distribution for the v_i's is to choose increment vectors w_i so that $v_i = (v_0 + \cdots + v_{i-1}) + w_i$, where w_i are independently and uniformly distributed on the $(n-1)$-dimensional hemisphere $S_+^{n-1} = \{x \in$

$\mathbf{R}^n \mid ||x|| = 1, x \cdot u > 0\}$, which consists of all the unit vectors in $\mathbf{R}^n$ in the u direction. We will call this distribution F.

Let $s_i = v_0 + \cdots + v_i$, $0 \le i \le m$. Then $v_0 = w_0$, $v_i = s_{i-1} + w_i$, and $s_i = 2^i w_0 + \cdots 2^0 w_i$ by an easy induction. We need some preliminaries. Let β_n denote the n-dimensional volume of the unit n-ball, let δ_{n-1} denote the $(n-1)$-dimensional volume of the unit $(n-1)$-sphere, then

$$\beta_n = \int_0^1 \delta_{n-1} r^{n-1} dr = \frac{\delta_{n-1}}{n}, \quad n \ge 2.$$

And

$$\beta_n = \int_{-1}^1 \beta_{n-1} (\sqrt{1-h^2})^{n-1} dh = 2\beta_{n-1} I_n,$$

where the integral $I_n = \int_0^{\frac{\pi}{2}} \sin^n \theta d\theta = \frac{n-1}{n} I_{n-2} = \ldots = \frac{\sqrt{\pi}}{2} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)}$. $I_n \approx \sqrt{\frac{\pi}{2n}}$ asymptotically for large n. Also

$$\beta_n = \int_0^{2\pi} \int_0^1 \beta_{n-2} (\sqrt{1-h^2})^{n-2} r dr d\theta = \beta_{n-2} \frac{2\pi}{n} = \frac{\pi^{n/2}}{\Gamma\left(\frac{n+2}{2}\right)}.$$

We will use the uniform distribution U on sets such as the hemisphere S_+^{n-1}, namely the Lebesgue measure on S_+^{n-1}, and we will denote a random variable X uniformly distributed on such a set S by $X \in_U S$. The following analysis is carried out using the exact Lebesgue measure. In the actual cryptographic protocols, this must be replaced by an exponentially close approximation on the set of rational points with exponentially large (polynomially bounded in binary length) denominators. The errors are exponentially small and thus insignificant. For clarity of presentation, we will state all results in terms of the exact Lebesgue measure.

Lemma 1. *Let $w, w' \in_U S_+^{n-1}$ be independently and uniformly distributed on the unit (northern) hemisphere. Let u be the north pole. Then the expectation of the inner product*

$$\mathbf{E}[w \cdot u] = \frac{1}{(n-1)I_{n-2}} = \frac{2}{(n-1)\sqrt{\pi}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \approx \sqrt{\frac{2}{\pi n}}.$$

Also,

$$\mathbf{E}[(w \cdot u)^2] = \frac{1}{n}.$$

$$\mathbf{E}[(w \cdot u)(w' \cdot u)] = (\mathbf{E}[w \cdot u])^2 \approx \frac{2}{\pi n}.$$

Proof For $w \in_U S_+^{n-1}$, the density function for the value of the inner product $h = w \cdot u$ is

$$p_{n-1}(h) = (\sqrt{1-h^2})^{n-3}/I_{n-2}.$$

Hence

$$\mathbf{E}[w \cdot u] = \int_0^1 h p_{n-1}(h) dh = \frac{1}{(n-1)I_{n-2}} = \frac{2}{(n-1)\sqrt{\pi}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \approx \sqrt{\frac{2}{\pi n}}.$$

Similarly

$$\mathbf{E}[(w \cdot u)^2] = \int_0^1 h^2 p_{n-1}(h) dh = \frac{1}{n}.$$

We note in passing that $\mathbf{E}_{S^{n-1}}[(w \cdot u)^2]$ over the whole unit sphere S^{n-1} is $1/n$ as well, by symmetry $h \to -h$.

The last equality follows from independence of w and w',

$$\mathbf{E}[(w \cdot u)(w' \cdot u)] = \mathbf{E}[hh'] = \mathbf{E}[h]\mathbf{E}[h'] = (\mathbf{E}[h])^2 \approx \frac{2}{\pi n}.$$

□

Lemma 2. *Let $w, w', w'' \in_U S_+^{n-1}$ be independently and uniformly distributed on the unit hemisphere. Then*

$$\mathbf{E}[w \cdot w'] = (\mathbf{E}[w \cdot u])^2 \approx \frac{2}{\pi n}.$$

$$\mathbf{E}[(w \cdot w')^2] = \frac{1}{n}.$$

$$\mathbf{E}[(w \cdot w')(w \cdot w'')] = (\mathbf{E}[w \cdot u])^2 \mathbf{E}[(w \cdot u)^2] \approx \frac{2}{\pi n^2}.$$

Proof Let $w, w', w'' \in_U S_+^{n-1}$. Choose a coordinate system so that u is the nth-coordinate. Then $w \cdot w' = \sum_{i=1}^n x_i(w) y_i(w')$. By linearity and independence $\mathbf{E}[w \cdot w'] = \sum_{i=1}^n \mathbf{E}[x_i]\mathbf{E}[y_i]$. For $i < n$, the symmetry $x_i \to -x_i$ implies that $\mathbf{E}[x_i] = 0$. For $i = n$, $x_n(w) = w \cdot u$, and similarly for $y_n(w')$. Then it follows that $\mathbf{E}[x_n] = \mathbf{E}[y_n] = \mathbf{E}[h]$, and

$$\mathbf{E}[w \cdot w'] = (\mathbf{E}[h])^2 \approx \frac{2}{\pi n}.$$

For $\mathbf{E}[(w \cdot w')^2]$, expand $(\sum_{i=1}^n x_i y_i)^2 = \sum_{i=1}^n x_i^2 y_i^2 + \sum_{1 \le i \ne j \le n} x_i y_i x_j y_j$. For $i \ne j$, at least one of i or j is not n, and by independence and the symmetry

$x_i \to -x_i$, we have $\mathbf{E}[x_i y_i x_j y_j] = \mathbf{E}[x_i x_j]\mathbf{E}[y_i y_j] = 0$. Thus the expectation of the second term is 0. By linearity and independence

$$\mathbf{E}[(w \cdot w')^2] = \sum_{i=1}^{n} \mathbf{E}[x_i^2]\mathbf{E}[y_i^2].$$

For $i = n$, it is $(\mathbf{E}[h^2])^2 = 1/n^2$. For $i < n$, by the symmetry $x_n \to -x_n$, it can be seen that $\mathbf{E}[x_i^2]$ is the same if we were to evaluate this expectation on the uniform distribution on the whole unit sphere. But on the whole sphere this is the same as $\mathbf{E}_{S^{n-1}}[x_n^2] = \mathbf{E}_{S^{n-1}}[h^2]$. This, however, by the symmetry $h \to -h$, is the same again if we were to evaluate it back on the hemisphere S_+^{n-1}. Hence ultimately $\mathbf{E}[x_i^2] = \mathbf{E}[h^2] = 1/n$, and $\mathbf{E}[x_i^2]\mathbf{E}[y_i^2] = 1/n^2$. It follows that

$$\mathbf{E}[(w \cdot w')^2] = 1/n.$$

Finally for $\mathbf{E}[(w \cdot w')(w \cdot w'')]$, we expand the product $(\sum_{i=1}^n x_i y_i)(\sum_{j=1}^n x_j z_j)$ $= \sum_{i=1}^n x_i^2 y_i z_i + \sum_{1 \le i \ne j \le n} x_i y_i x_j z_j$. For $i \ne j$, at least one of them is not n, so that either $\mathbf{E}[y_i] = 0$ or $\mathbf{E}[z_j] = 0$, thus $\sum_{1 \le i \ne j \le n} \mathbf{E}[x_i^2]\mathbf{E}[y_i]\mathbf{E}[z_j] = 0$. Then

$$\mathbf{E}[(w \cdot w')(w \cdot w'')] = \sum_{i=1}^{n} \mathbf{E}[x_i^2]\mathbf{E}[y_i]\mathbf{E}[z_i].$$

For $i < n$, $\mathbf{E}[y_i] = 0$ by symmetry as before. For $i = n$, it is $\mathbf{E}[h^2](\mathbf{E}[h])^2 \approx \frac{2}{\pi n^2}$. □

For the distribution F, we will show that the secret information u is not safe. In fact we claim that s_m can be used to approximate the direction u. Consider $s_m \cdot u = 2^m(w_0 \cdot u) + \cdots + 2^0(w_m \cdot u)$.

$$\mathbf{E}_F[s_m \cdot u] = (2^m + \cdots + 2^0)\mathbf{E}[w \cdot u] \approx 2^{m+1}\sqrt{\frac{2}{\pi n}}.$$

Next we compute the variance $\mathbf{Var}_F[s_m \cdot u]$. First $(s_m \cdot u)^2 = 2^{2m}(w_0 \cdot u)^2 + \cdots + 2^0(w_m \cdot u)^2 + \sum_{0 \le i \ne j \le m} 2^{m-i}2^{m-j}(w_i \cdot u)(w_j \cdot u)$. So

$$\begin{aligned}
\mathbf{E}_F[(s_m \cdot u)^2] &= (2^{2m} + \cdots + 2^0)\mathbf{E}[(w \cdot u)^2] + \sum_{0 \le i \ne j \le m} 2^{i+j}\mathbf{E}[(w \cdot u)(w' \cdot u)] \\
&= \frac{4^{m+1} - 1}{3n} + \left[\sum_{0 \le i,j \le m} 2^{i+j} - \sum_{0 \le i \le m} 2^{2i}\right](\mathbf{E}[w \cdot u])^2 \\
&\approx \frac{4^{m+1}}{3n} + \frac{4^{m+2}}{3\pi n} = \frac{4^{m+1}}{3\pi n}(4 + \pi).
\end{aligned}$$

It follows that

$$\mathbf{Var}_F[s_m \cdot u] \approx \frac{4^{m+1}}{3\pi n}(\pi - 2).$$

We note that the normalized ratio

$$\frac{\mathbf{E}_F[s_m \cdot u]}{\sqrt{\mathbf{Var}_F[s_m \cdot u]}} \approx \sqrt{\frac{6}{\pi - 2}} \approx 2.2925564.$$

This indicates that s_m has a significant correlation with the hidden direction u, and hence u can not be considered secure under this distribution F.

More directly it can be shown that

$$\frac{\mathbf{E}_F[s_m \cdot (v_m - v_{m-1})]}{\mathbf{E}_F[||s_m||]\mathbf{E}_F[||v_m - v_{m-1}||]} \geq 1$$

asymptotically. Thus one can expect s_m to be used to distinguish v_m from the others.

We now return to our chosen distribution D, and show that in this distribution there is no easy statistical leakage. In this distribution, $v_i = \frac{2^i}{M}u + \sqrt{1 - \frac{2^{2i}}{M^2}}\rho_i$, and ρ_i are independently and uniformly distributed on the $(n-2)$-dimensional unit sphere orthogonal to u. Recall $||v_i|| = 1$. Let $s'_i = v_0 + \ldots + v_i$. We consider $||s'_m||^2$ and $s'_m \cdot u$. Clearly $||s'_i||^2 = (m+1) + \sum_{0 \leq i \neq j \leq m}(v_i \cdot v_j)$.

Lemma 3. *For* $0 \leq i \neq j \leq m$,

$$\mathbf{E}_D[v_i \cdot v_j] = \frac{2^{i+j}}{M^2}, \quad \textit{and} \quad \mathbf{Var}_D[v_i \cdot v_j] = \frac{1}{n-1}\left(1 - \frac{2^{2i}}{M^2}\right)\left(1 - \frac{2^{2j}}{M^2}\right).$$

Proof We have $v_i \cdot v_j = \frac{2^{i+j}}{M^2} + \sqrt{1 - \frac{2^{2i}}{M^2}}\sqrt{1 - \frac{2^{2j}}{M^2}}(\rho_i \cdot \rho_j)$. By symmetry, $\mathbf{E}_{S^{n-2}}[\rho_i \cdot \rho_j] = 0$, so that $\mathbf{E}_D[v_i \cdot v_j] = \frac{2^{i+j}}{M^2}$. Thus,

$$\mathbf{Var}_D[v_i \cdot v_j] = \left(1 - \frac{2^{2i}}{M^2}\right)\left(1 - \frac{2^{2j}}{M^2}\right)\mathbf{Var}_D[\rho_i \cdot \rho_j].$$

We have $\mathbf{Var}_D[\rho_i \cdot \rho_j] = \mathbf{E}_D[(\rho_i \cdot \rho_j)^2] = \int_{-1}^{1} h^2(p_{n-2}(h)/2)dh = \frac{1}{n-1}$. The lemma follows. □

Now

$$\mathbf{E}_D[||s'_m||^2] = (m+1) + \sum_{0 \leq i \neq j \leq m} \frac{2^{i+j}}{M^2},$$

and $\sum_{0 \leq i \neq j \leq m} 2^{i+j} = \sum_{0 \leq i,j \leq m} 2^{i+j} - \sum_{i=0}^{m} 2^{2i} \approx \frac{8}{3}(2^m/M)^2$. Hence $\mathbf{E}_D[||s'_m||^2] \approx (m+1) + \frac{2^{2m+3}}{3}$. Ignoring the exponentially small term $2^{2m}/M^2$, $\mathbf{E}_D[||s'_m||^2] \approx m+1$.

One should compare this with the uniform distribution U for which all v_i's are independently and uniformly distributed on S^{n-1}. In this case $\mathbf{E}_U[v_i \cdot v_j] = 0$, for $i \neq j$, and $\mathbf{E}_U[||s'_m||^2] = m+1$.

We next evaluate the variance $\mathbf{Var}_D[||s'_m||^2]$.

$$\begin{aligned}\mathbf{Var}_D[||s'_m||^2] &= \mathbf{E}_D\left[\left\{\sum_{0\le i\ne j\le m}(v_i\cdot v_j - E_D[v_i\cdot v_j])\right\}^2\right]\\ &= 4\mathbf{E}_D\left[\left\{\sum_{0\le i<j\le m}\sqrt{1-\frac{2^{2i}}{M^2}}\sqrt{1-\frac{2^{2j}}{M^2}}(\rho_i\cdot\rho_j)\right\}^2\right]\\ &= 4\mathbf{E}_D\left[\sum_{(i<j)}\left(1-\frac{2^{2i}}{M^2}\right)\left(1-\frac{2^{2j}}{M^2}\right)(\rho_i\cdot\rho_j)^2\right]\\ &+ 4\mathbf{E}_D\left[\sum_{(i<j)\ne(i'<j')}c_{(ij)}c_{(i'j')}(\rho_i\cdot\rho_j)(\rho_{i'}\cdot\rho_{j'})\right].\end{aligned}$$

For $(i<j)\ne(i'<j')$, there are two cases. If (i,j,i',j') are all distinct indices, then clearly $\rho_i\cdot\rho_j$ and $\rho_{i'}\cdot\rho_{j'}$ are independent. Thus $\mathbf{E}_D[(\rho_i\cdot\rho_j)(\rho_{i'}\cdot\rho_{j'})] = \mathbf{E}_D[\rho_i\cdot\rho_j]\mathbf{E}_D[\rho_{i'}\cdot\rho_{j'}] = 0$. If there are only 3 distinct indices among (i,j,i',j'), say $i=i'$, then by fixing ρ_i, the conditional distribution of $\rho_i\cdot\rho_j$ and $\rho_i\cdot\rho_{j'}$ over ρ_j and $\rho_{j'}$ are independent, and $\mathbf{E}_D[(\rho_i\cdot\rho_j)|\rho_i] = \mathbf{E}_D[(\rho_i\cdot\rho_{j'})|\rho_i] = 0$. Thus in any case $\mathbf{E}_D[(\rho_i\cdot\rho_j)(\rho_{i'}\cdot\rho_{j'})] = 0$, for $(i<j)\ne(i'<j')$, and

$$\mathbf{Var}_D[||s'_m||^2] = 4\sum_{0\le i<j\le m}\left(1-\frac{2^{2i}}{M^2}\right)\left(1-\frac{2^{2j}}{M^2}\right)\mathbf{E}_D\left[(\rho_i\cdot\rho_j)^2\right].$$

We have $\mathbf{E}_D\left[(\rho_i\cdot\rho_j)^2\right] = \mathbf{E}_{S^{n-2}}[h^2] = 1/(n-1)$. Ignoring exponentially small terms such as $2^m/M$, we get

$$\mathbf{Var}_D[||s'_m||^2] \approx \frac{4}{n-1}\binom{m+1}{2}.$$

This is to be compared to the uniform distribution U. Again $||s'_i||^2 = (m+1) + \sum_{0\le i\ne j\le m}(v_i\cdot v_j)$. But for the uniform distribution U, $\mathbf{E}_U[v_i\cdot v_j] = 0$, for $i\ne j$, and so

$$\begin{aligned}\mathbf{Var}_U[||s'_m||^2] &= \mathbf{E}_U\left[\left\{\sum_{0\le i\ne j\le m}(v_i\cdot v_j)\right\}^2\right]\\ &= 4\mathbf{E}_U\left[\sum_{0\le i<j\le m}(v_i\cdot v_j)^2 + \sum_{(i<j)\ne(i'<j')}(v_i\cdot v_j)(v_{i'}\cdot v_{j'})\right].\end{aligned}$$

By the same argument, $E_U[(v_i\cdot v_j)(v_{i'}\cdot v_{j'})] = 0$, for all $(i<j)\ne(i'<j')$. Hence $\mathbf{Var}_U[||s'_m||^2] = 4\sum_{0\le i<j\le m}\mathbf{E}_U[(v_i\cdot v_j)^2]$, where $\mathbf{E}_U[(v_i\cdot v_j)^2]$ is $\mathbf{E}[h^2]$

over $(n-1)$-dimensional unit sphere, and thus equal to $1/n$ (see the proof of Lemma 1). It follows that

$$\mathbf{Var}_U[||s'_m||^2] = \frac{4}{n}\binom{m+1}{2}.$$

We conclude that at least in terms of the length of the sum $||s_m|| = ||v_0 + \ldots + v_m||$, our distribution D behaves very much like the uniform distribution U.

We return to the correlation between s'_m and u. It is easy to see that with distribution D

$$s'_m \cdot u = \sum_{i=0}^{m} \frac{2^i}{M} \approx \frac{2^{m+1}}{M},$$

which is exponentially small. Also since it is a constant $\mathbf{Var}_D[s'_m \cdot u] = 0$. For the uniform distribution U,

$$s'_m \cdot u = \sum_{i=0}^{m} v_i \cdot u,$$

and $\mathbf{E}_U[s'_m \cdot u] = 0$. For the variance

$$\begin{aligned}
\mathbf{Var}_U[s'_m \cdot u] &= \mathbf{E}_U[(s'_m \cdot u)^2] \\
&= \mathbf{E}_U\left[\left(\sum_{i=0}^{m}(v_i \cdot u)\right)^2\right] \\
&= \mathbf{E}_U\left[\sum_{i=0}^{m}(v_i \cdot u)^2 + \sum_{0\le i\ne j\le m}(v_i \cdot u)(v_j \cdot u)\right].
\end{aligned}$$

By independence $\mathbf{E}_U[(v_i \cdot u)(v_j \cdot u)] = \mathbf{E}_U[v_i \cdot u]\mathbf{E}_U[v_j \cdot u] = 0$ for $i \ne j$. Also $\mathbf{E}_U[(v_i \cdot u)^2] = 1/n$. Hence $\mathbf{Var}_U[s'_m \cdot u] = m/n$. Therefore statistically one can not deduce much from $s'_m \cdot u$ in the distribution D, since it is exponentially small, and well within the range in which this value would have been under the uniform distribution, where $\mathbf{E}_U[s'_m \cdot u] = 0$ and $\mathbf{Var}_U[s'_m \cdot u] = \Omega(1)$.

In fact, suppose $u' \in S^{n-1}$ is any unit vector, $u' \ne \pm u$. The estimates of $\mathbf{E}_U[s'_m \cdot u'] = 0$ and $\mathbf{Var}_U[s'_m \cdot u'] = m/n = \Omega(1)$ are still valid. Let Π be the 2-dimensional plane spanned by u and u'. Let $u' = (\cos\theta)u + (\sin\theta)u^\perp$, where the unit vector $u^\perp \perp u$. Then we can choose a coordinate system such that $u^\perp$ is the $(n-1)$st coordinate for ρ_i, and $\mathbf{E}_D[\rho_i \cdot u^\perp] = \mathbf{E}_{S^{n-2}}[h] = 0$. Therefore $\mathbf{E}_D[s'_m \cdot u'] = (\cos\theta)\mathbf{E}_D[s'_m \cdot u]$. Thus $|\mathbf{E}_D[s'_m \cdot u']| \le 2^{m+1}/M$, which is exponentially small. This implies that s'_m has correlation with no particular direction, the same as under the uniform distribution U.

References

1. M. Ajtai. Generating hard instances of lattice problems. In *Proc. 28th Annual ACM Symposium on the Theory of Computing*, 1996. Full version available from ECCC, *Electronic Colloquium on Computational Complexity* TR96-007, at `http://www.eccc.uni-trier.de/eccc/`.
2. M. Ajtai. The shortest vector problem in L_2 is NP-hard for randomized reductions. *Electronic Colloquium on Computational Complexity*, TR97-047 at `http://www.eccc.uni-trier.de/eccc/`.
3. M. Ajtai and C. Dwork. A public-key cryptosystem with worst-case/average-case equivalence. 1996. Available from ECCC, *Electronic Colloquium on Computational Complexity* TR96-065, at `http://www.eccc.uni-trier.de/eccc/`.
4. S. Arora, L. Babai, J. Stern, and Z. Sweedyk. The hardness of approximate optima in lattices, codes, and systems of linear equations. In *Proc. 34th IEEE Symposium on Foundations of Computer Science (FOCS)*, 1993, 724-733.
5. J-Y. Cai and A. Nerurkar. An Improved Worst-Case to Average-Case Connection for Lattice Problems. In *Proc. 38th IEEE Symposium on Foundations of Computer Science (FOCS)*, 1997, 468–477.
6. J-Y. Cai and A. Nerurkar. Approximating the SVP to within a factor $\left(1 + \frac{1}{\dim^{\epsilon}}\right)$ is NP-hard under randomized reductions. Available from ECCC, *Electronic Colloquium on Computational Complexity* TR97-059, at `http://www.eccc.uni-trier.de/eccc/`.
7. J-Y. Cai. A Primal-Dual Relation for Lattices and the Complexity of Shortest Lattice Vector Problem. To appear in *Theoretical Computer Science*.
8. J-Y. Cai. A new transference theorem and applications to Ajtai's connection factor. *Electronic Colloquium on Computational Complexity* TR98-005, at `http://www.eccc.uni-trier.de/eccc/`.
9. O. Goldreich and S. Goldwasser. On the Limits of Non-Approximability of Lattice Problems. *Electronic Colloquium on Computational Complexity* TR97-031, at `http://www.eccc.uni-trier.de/eccc/`.
10. O. Goldreich, S. Goldwasser, and S. Halevi. Collision-free hashing from lattice problems. 1996. Available from ECCC, *Electronic Colloquium on Computational Complexity* TR96-042, at `http://www.eccc.uni-trier.de/eccc/`.
11. O. Goldreich, S. Goldwasser, and S. Halevi. Public-key cryptosystems from lattice reduction problems. 1996. Available from ECCC, *Electronic Colloquium on Computational Complexity* TR96-056, at `http://www.eccc.uni-trier.de/eccc/`.
12. J. C. Lagarias. The computational complexity of simultaneous diophantine approximation problems. *SIAM Journal of Computing*, Volume 14, page 196–209, 1985.
13. J. C. Lagarias, H. W. Lenstra, and C. P. Schnorr. Korkin-Zolotarev Bases and Successive Minima of a Lattice and its Reciprocal Lattice. *Combinatorica*, 10:(4), 1990, 333-348.
14. A. K. Lenstra, H. W. Lenstra, and L. Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261:515–534, 1982.
15. A.M. Odlyzko. The rise and fall of knapsack cryptosystems. in *Cryptology and Computational Number Theory*, American Mathematical Society, pp. 75-88, 1990.
16. C. P. Schnorr. A hierarchy of polynomial time basis reduction algorithms. *Theory of Algorithms*, pages 375–386, 1985.
17. P. van Emde Boas. Another NP-complete partition problem and the complexity of computing short vectors in lattices. Technical Report 81-04, Mathematics Department, University of Amsterdam, 1981.