Solving Random Subset Sum Problem by l_p-norm SVP Oracle[*]

Gengran Hu, Yanbin Pan, and Feng Zhang

Key Laboratory of Mathematics Mechanization, NCMIS,
Academy of Mathematics and Systems Science, Chinese Academy of Sciences
Beijing 100190, China
hudiran10@mails.ucas.ac.cn, {panyanbin,zhangfeng}@amss.ac.cn

Abstract. It is well known that almost all random subset sum instances with density less than 0.6463... can be solved with an l_2-norm SVP oracle by Lagarias and Odlyzko. Later, Coster *et al.* improved the bound to 0.9408... by using a different lattice. In this paper, we generalize this classical result to l_p-norm. More precisely, we show that for $p \in \mathbb{Z}^+$, an l_p-norm SVP oracle can be used to solve almost all random subset sum instances with density bounded by δ_p, where $\delta_1 = 0.5761$ and $\delta_p = 1/(\frac{1}{2^p}\log_2(2^{p+1}-2) + \log_2(1 + \frac{1}{(2^p-1)(1-(\frac{1}{2^{p+1}-2})^{(2^p-1)})}))$ for $p \geq 3$(asymptotically, $\delta_p \approx 2^p/(p+2)$). Since δ_p goes increasingly to infinity when p tends to infinity, it can be concluded that an l_p-norm SVP oracle with bigger p can solve more subset sum instances. An interesting phenomenon is that an l_p-norm SVP oracle with $p \geq 3$ can help solve almost all random subset sum instances with density one, which are thought to be the most difficult instances.

Keywords: SVP, random subset sum problems, lattice, l_p-norm.

1 Introduction

Lattices are discrete subgroup in $\mathbb{R}^n$ and have many important applications in both cryptanalysis and cryptographic constructions. Many lattice-based cryptographic primitives have been presented, such as the public-key cryptosystems [1,2,21,9,11], the digital signature scheme NTRUSign [12] and the fully homomorphic encryption [8]. Usually, the securities of these schemes can be based on the hardness of some lattice problems, like SVP (the shortest vector problem). SVP refers to finding a shortest non-zero vector in a given lattice and is one of the most famous computational problems of lattice. Many famous algorithms are proposed to solve SVP, including the famous LLL algorithm [14]. These algorithms can also be used to attack knapsack-based public-key cryptosystems (See [15] for more details).

[*] This work was supported in part by the NNSF of China (No.11071285, No.11201458, and No.61121062), in part by 973 Project (No. 2011CB302401) and in part by the National Center for Mathematics and Interdisciplinary Sciences, CAS.

H. Krawczyk (Ed.): PKC 2014, LNCS 8383, pp. 399–410, 2014.

The knapsack problem, or the subset sum problem (SSP), is a well-known NP-hard problem. It asks to choose some elements in a given set such that the sum of these elements is exactly equal to a given number. When all of the elements of the set are uniformly random over some set, it comes to the random subset sum problem (RSSP), which is also a significant computational problem.

The hardness of RSSP is still not clear. However, it seems that there is a very close relationship between the hardness of RSSP and its density. When the density is large enough, it can be solved via dynamic programming. When the density is small enough, it can be solved by LLL algorithm [15]. In [13], Impagliazzo and Naor showed that the hardest instances of RSSP lie in those with density equal to 1.

Some relations between SVP and RSSP have been exploited. In 1985, Lagarias and Odlyzko [15] showed that the l_2-norm SVP oracle can be used to solve almost all random subset sum instances with density bounded by 0.6463 when the size of the subset sum instance is large enough. Later, Coster *et al.* [5] improved this bound to 0.9408. However, it is a long standing open problem to solve the RSSP instances with density 1 using the lattice l_2-norm SVP oracle.

In this work, we give a very interesting result that any l_p-norm SVP oracle ($p > 2$) can help to solve the RSSP with density 1 efficiently. More precisely, if $p \in \mathbb{Z}^+$, an l_p-norm SVP oracle can be used to solve almost all random subset sum instances with density bounded by $\delta_p = 1/(\frac{1}{2^p}\log_2(2^{p+1}-2) + \log_2(1 + \frac{1}{(2^p-1)(1-(\frac{1}{2^{p+1}-2})^{(2^p-1)})})))$. It is easy to see that δ_p goes increasingly to infinity as p tends to infinity, which implies that an l_p-norm SVP oracle with bigger p can solve more subset sum instances. Especially, an l_∞-norm SVP oracle can solve all the subset sum instances, which coincides with the deterministic reduction from subset sum problem to l_∞-norm SVP. It seems that the hardness of l_p-norm SVP increases as p gets bigger. However, in practice, the existing SVP algorithms are mostly in l_2-norm, even the l_p-norm SVP algorithm in [6] uses the MV algorithm(an l_2-norm SVP algorithm in [18]) as a starting point.

In fact, it is well known that the l_∞-norm SVP is NP-hard under deterministic reduction, whereas SVP for other norms are proved to be NP-hard under only probabilistic reductions (see [3,17,19]). In addition, Regev and Rosen [22] proved for any $\epsilon > 0$, l_2-norm SVP$_{1+\epsilon}$ can be probabilistically reduced to l_p-norm SVP for all $1 \leq p \leq \infty$ which showed that the l_2-norm SVP$_{1+\epsilon}$ is easiest. Unfortunately, reduction from exact l_2-norm SVP to l_p-norm SVP has still not been found.

We would like to point out that if RSSP can be proved to be NP-hard, then by our result, we can prove l_p-norm SVP ($p > 2, p \in \mathbb{Z}^+$) is NP-hard under probabilistic reduction. Such a reduction will be more simple and clear, compared to the previous reductions.

Moreover, as a byproduct, we give an upper bound of the number of the integer points in an l_p ball, which is shown to be very nice for $p \geq 3$.

Roadmap. The remainder of the paper is organized as follows. In Section 2, we give some preliminaries needed. In Section 3, we describe our probabilistic reduction from random subset sum problem to l_p-norm SVP in details. Finally, we give a short conclusion in Section 4.

2 Preliminaries

We denote by $\mathbb{Z}$ the integer ring. We use bold letters to denote vectors. If $\mathbf{v} \in \mathbb{R}^n$ is a vector, then we denote by v_i the i-th entry of $\mathbf{v}$. Let $\|\mathbf{v}\|_p$ be the l_p norm of $\mathbf{v}$, that is, $\|\mathbf{v}\|_p = (\sum_{i=1}^n |v_i|^p)^{1/p}$.

2.1 Lattice

Given a matrix $B = (b_{ij}) \in \mathbb{R}^{m\times n}$ with rank n, the lattice $\mathcal{L}(B)$ spanned by the columns of B is

$$\mathcal{L}(B) = \{Bx = \sum_{i=1}^{n} x_i b_i | x_i \in \mathbb{Z}\},$$

where b_i is the i-th column of B. We call m the dimension of $\mathcal{L}(B)$ and n its rank.

Definition 1 (l_p-norm SVP). *Given a lattice basis B, the l_p-norm SVP asks to find a nonzero vector in $\mathcal{L}(B)$ with the smallest l_p-norm.*

2.2 Random Subset Sum Problem

Given $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ distributed uniformly in $[1, A]^n$ and $s = \sum_{i=1}^n e_i a_i$ where $\mathbf{e} = (e_1, e_2, \ldots, e_n) \in \{0,1\}^n$, RSSP refers to finding some $\mathbf{c} = (c_1, c_2, \ldots, c_n) \in \{0,1\}^n$ such that $s = \sum_{i=1}^n c_i a_i$ without knowing $\mathbf{e}$. Notice that the solution $\mathbf{c}$ may not be the original $\mathbf{e}$.

The density of these a_i's is defined by

$$d = \frac{n}{\log_2(A)}.$$

It was shown by Lagarias and Odlyzko [15] that almost all the subset sum problem with density less than $0.6463\ldots$ would be solved in polynomial time with a single call to an oracle that can find the shortest vector in a special lattice. Later, Coster *et al.* [5] improved the bound to $0.9408\ldots$ by finding a shortest nonzero vector with an l_2-norm SVP oracle in the following lattice spanned by the columns of

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & \frac{1}{2} \\ 0 & 1 & \ldots & 0 & \frac{1}{2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & \frac{1}{2} \\ Na_1 & Na_2 & \ldots & Na_n & Ns \end{pmatrix},$$

where N is a big enough integer.

2.3 Estimation of the Combinatorial Number

By Stirling's Formula, we have the following estimation for the combinatorial number,

$$\binom{\alpha n}{\beta n} = \widetilde{O}(2^{\alpha H(\beta/\alpha)n}),$$

where
$H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ and $\widetilde{O}(f(n)) = O(f(n) * poly(\log(f(n))))$.

3 Solving Random Subset Sum Problem by l_p-norm SVP Oracle

3.1 The Upper Bound of the Number of Integer Points in an l_p-Ball

We first give some results about the number of the integer points in an l_p-ball, that is, $\#\{\mathbf{x} \in \mathbb{Z}^{n+1} | \|\mathbf{x}\|_p \le \frac{1}{2}(n+1)^{\frac{1}{p}}\}$.

Theorem 1. *For all $n \ge 1$,*

- *If $p = 1$ and n large enough,*

$$\#\{\mathbf{x} \in \mathbb{Z}^n | \|\mathbf{x}\|_1 \le \frac{1}{2}n\} \le 2^{c_1 n},$$

  where $c_1 = 1.7357$.
- *If $p = 2$,*

$$\#\{\mathbf{x} \in \mathbb{Z}^n | \|\mathbf{x}\|_2 \le \frac{1}{2}\sqrt{n}\} \le 2^{c_2 n},$$

  where $c_2 = 1.0628$.
- *If $p \ge 3$ and $p \in \mathbb{Z}^+$,*

$$\#\{\mathbf{x} \in \mathbb{Z}^n | \|\mathbf{x}\|_p \le \frac{1}{2}n^{\frac{1}{p}}\} \le 2^{c_p n},$$

  where $c_p \approx \frac{1}{2^p}\log_2(2^{p+1}-2) + \log_2(1 + \frac{1}{2^p - 1})$.

Proof. We will prove the theorem in three cases.

- $p = 1$:
  For simplicity, we assume n is even (the case when n is odd is similar). Let $R(m,n) \triangleq \#\{\mathbf{x} \in \mathbb{Z}^n, \mathbf{x}$ has m nonzero entries $| \quad \|\mathbf{x}\|_1 \le \frac{1}{2}n\}$, then

$$\#\{\mathbf{x} \in \mathbb{Z}^n | \|\mathbf{x}\|_1 \le \frac{1}{2}n\} = \sum_{m=0}^{n/2} R(m,n).$$

It is easy to know that $R(m,n) = 2^m\binom{n}{m}\sum_{j=m}^{n/2}\binom{j-1}{m-1} = 2^m\binom{n}{m}\binom{n/2}{m}$. Assume $R(m_n,n) = \max\limits_{m} R(m,n)$, then $\#\{\mathbf{x}\in\mathbb{Z}^n \,|\, \|\mathbf{x}\|_1 \le \frac{1}{2}n\} \le \frac{1}{2}n(R(m_n,n))$. Noticing that

$$\begin{aligned} R(m+1,n)/R(m,n) &= \frac{2^{m+1}\binom{n}{m+1}\binom{n/2}{m+1}}{2^m\binom{n}{m}\binom{n/2}{m}} \\ &= \frac{(n-m)(n-2m)}{(m+1)^2} \end{aligned}$$

is decreasing with respect to m, we have

$$\begin{cases} R(m_n,n)/R(m_n-1,n) \ge 1, \\ R(m_n+1,n)/R(m_n,n) \le 1, \end{cases}$$

which implies that

$$\begin{cases} m_n \le 0.381966n + 0.658359, \\ m_n \ge 0.381966n - 0.828427, \end{cases}$$

since $m_n \le n/2$. We obtain $m_n \approx 0.381966n$. Thus, we have the bound

$$\#\{\mathbf{x}\in\mathbb{Z}^n \,|\, \|\mathbf{x}\|_1 \le \frac{1}{2}n\} \le \frac{1}{2}n2^{0.381966n}\binom{0.5n-1}{0.381966n-1}\binom{n}{0.381966n}.$$

Using the estimation $\binom{\alpha n}{\beta n} = \widetilde{O}(2^{\alpha H(\beta/\alpha)n})$, finally we have for n large enough

$$\#\{\mathbf{x}\in\mathbb{Z}^n \,|\, \|\mathbf{x}\|_1 \le \frac{1}{2}n\} = \widetilde{O}(2^{0.381966n+0.39422n+0.9594187n}) = \widetilde{O}(2^{1.7356047n}) \le 2^{1.7357n}.$$

- If $p = 2$:
  It has been proven in Section 3 in [5].
- If $p \ge 3$ and $p \in \mathbb{Z}^+$:
  Let $\theta(z) = 1 + 2\sum_{i=1}^{\infty} z^{i^p}$ and $r_n(k)$ be the number of integer solutions to
  $$\sum_{i=1}^{n} |x_i|^p = k.$$
  Then
  $$(\theta(z))^n = \sum_{k=0}^{\infty} r_n(k)z^k.$$

For all $x > 0$, we have

$$\begin{aligned}\#\{\mathbf{x} \in \mathbb{Z}^n | \|\mathbf{x}\|_p \le \frac{1}{2} n^{\frac{1}{p}}\} &= \#\{\mathbf{x} \in \mathbb{Z}^n | \|\mathbf{x}\|_p^p \le \frac{1}{2^p} n\} \\ &= \sum_{k \le \frac{1}{2^p} n} r_n(k) \\ &\le \sum_{k \le \frac{1}{2^p} n} r_n(k) e^{\frac{1}{2^p} nx} e^{-kx} \\ &\le \sum_{k=0}^{\infty} r_n(k) e^{\frac{1}{2^p} nx} e^{-kx} \\ &= e^{\frac{1}{2^p} nx} \sum_{k=0}^{\infty} r_n(k) e^{-kx} \\ &= e^{\frac{1}{2^p} nx} (\theta(e^{-x}))^n.\end{aligned}$$

Let

$$f_p(x) = \frac{1}{2^p} x + \ln \theta(e^{-x}).$$

We have

$$\#\{\mathbf{x} \in \mathbb{Z}^n | \|\mathbf{x}\|_p \le \frac{1}{2} n^{\frac{1}{p}}\} \le e^{f_p(x) n} = 2^{(\log_2 e) f_p(x) n}$$

holds for all $x > 0$.
So we only need to compute $\min\limits_{x>0} f_p(x)$. It is difficult to give the exact value of $\min\limits_{x>0} f_p(x)$. Next we give an upper bound for $\min\limits_{x>0} f_p(x)$.

Noticing that

$$f_p(x) = \frac{1}{2^p} x + \ln \theta(e^{-x}) = \frac{1}{2^p} x + \ln(1 + 2e^{-1^p x} + 2e^{-2^p x} + 2e^{-3^p x} + \cdots + 2e^{-k^p x} + \cdots),$$

we define

$$l_p(x) \triangleq \frac{1}{2^p} x + \ln(1 + 2e^{-x})$$

and

$$\begin{aligned}u_p(x) &\triangleq \frac{1}{2^p} x + \ln(1 + 2e^{-x} + 2e^{-2^p x} + 2e^{-(2^p + 2^p - 1)x} + \cdots + 2e^{-((k-1)2^p - (k-2))x} + \cdots) \\ &= \frac{1}{2^p} x + \ln(1 + \frac{2e^{-x}}{1 - e^{-(2^p - 1)x}}).\end{aligned}$$

When $p \ge 1$, the difference sequence $(2^p - 1^p, 3^p - 2^p, 4^p - 3^p, \cdots)$ is not decreasing, then for $k \ge 2$,

$$\begin{aligned}2e^{-k^p x} &= 2e^{-x(1 + (2^p - 1^p) + (3^p - 2^p) + \cdots + (k^p - (k-1)^p))} \\ &\le 2e^{-x(1 + (2^p - 1^p) + (2^p - 1) + \cdots + (2^p - 1))} \\ &= 2e^{-((k-1)2^p - (k-2))x}\end{aligned}$$

So we have

$$l_p(x) \leq f_p(x) \leq u_p(x)$$

holds for all $x > 0$, which implies

$$\min_{x>0} l_p(x) \leq \min_{x>0} f_p(x) \leq \min_{x>0} u_p(x).$$

Because $l_p(x)$ takes the minimum

$$l_p(x_0(p)) = \frac{1}{2^p} \ln(2^{p+1} - 2) + \ln(1 + \frac{1}{2^p - 1})$$

at

$$x_0(p) = \ln(2^{p+1} - 2)$$

and

$$u_p(x_0(p)) = \frac{1}{2^p} \ln(2^{p+1} - 2) + \ln(1 + \frac{1}{(2^p - 1)(1 - (\frac{1}{2^{p+1}-2})^{(2^p-1)})}),$$

we have an interval estimate $[l_p(x_0(p)), \quad u_p(x_0(p))]$ for $\min_{x>0} f_p(x)$ since

$$l_p(x_0(p)) = \min_{x>0} l_p(x) \leq \min_{x>0} f_p(x) \leq \min_{x>0} u_p(x) \leq u_p(x_0(p)).$$

Taking $c_p = \log_2 e \cdot u_p(x_0(p))$, the result for $p \geq 3$ follows.

We would like to point out that $2^{c_p n}$ is a very nice estimation of the number of integer points in the l_p ball $\{\mathbf{x} \in \mathbb{Z}^n | \|\mathbf{x}\|_p \leq \frac{1}{2} n^{\frac{1}{p}}\}$ for $p \geq 3$. In fact, we can easily have an asymptotic rough lower bound for the number of integer points by just considering those vectors in the ball with exactly $\frac{1}{2^p} n$ entries in $\{-1, 1\}$ and other entries equal to 0. The total number of such vectors is $2^{\frac{1}{2^p} n} \cdot \binom{n}{\frac{1}{2^p} n}$, which is approximately equal to $2^{(H(\frac{1}{2^p}) + \frac{1}{2^p})n}$. Hence for $p \in \mathbb{Z}^+$ and n large enough, we have

$$\#\{\mathbf{x} \in \mathbb{Z}^n | \|\mathbf{x}\|_p \leq \frac{1}{2} n^{\frac{1}{p}}\} \geq 2^{k_p n},$$

where $k_p = \frac{p+1}{2^p} - (1 - \frac{1}{2^p}) \log_2(1 - \frac{1}{2^p})$. Interestingly, we find that k_p is exactly the total lower bound $\log_2 e \cdot l_p(x_0(p))$ obtained above.

The table below gives the values of $l_p(x_0(p))(= \ln 2 \cdot k_p)$ and $u_p(x_0(p))$ for p from three to ten.

p	3	4	5	6	7	8	9	10
l_p	0.4634	0.2771	0.1607	0.0913	0.0511	0.2827	0.0155	0.0084
u_p	0.4634	0.2771	0.1607	0.0913	0.0511	0.2827	0.0155	0.0084

It can be seen that $u_p(x_0(p))$ is a very good estimation of $\min_{x>0} f_p(x)$, since for $p \geq 3$, $l_p(x_0(p))$ and $u_p(x_0(p))$ are nearly the same. Similarly, $2^{c_p n}$ is a very nice estimation of the number of integer points in the l_p ball for $p \geq 3$, since the upper bound and the lower bound are also nearly the same. In fact, the asymptotic forms for $l_p(x_0(p))$ and $u_p(x_0(p))$ are the same:

$$l_p(x_0(p)) \approx \ln 2 \cdot \frac{p+2}{2^p}, u_p(x_0(p)) \approx \ln 2 \cdot \frac{p+2}{2^p}.$$

3.2 Solving Random Subset Sum Problem by l_p-norm SVP Oracle

To solve the subset sum problem defined by $a_i (1 \leq i \leq n)$ and s, we consider the lattice $\mathcal{L}(B)$ generated by the columns of B where

$$B = \begin{pmatrix} 1 & 0 & \dots & 0 & \frac{1}{2} \\ 0 & 1 & \dots & 0 & \frac{1}{2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & \frac{1}{2} \\ 0 & 0 & \dots & 0 & \frac{1}{2} \\ Na_1 & Na_2 & \dots & Na_n & Ns \end{pmatrix},$$

and $N > \frac{1}{2}(n+1)^{\frac{1}{p}}$ is an positive integer. Notice that our lattice is a little different from Coster *et al.*'s, which leads a more simple reduction. The additional row in the lattice basis matrix can bound the last integer coefficient more tightly(see section 3 of [5] for more details).

Any $\mathbf{x} = (x_1, x_2 \dots x_n, x_{n+1}, x_{n+2})^T \in \mathcal{L}(B)$ can be written as

$$\begin{cases} x_i = w_i + \frac{1}{2}w & (i = 1, 2 \dots n) \\ x_{n+1} = \frac{1}{2}w \\ x_{n+2} = N(\sum_{i=1}^{n} w_i a_i + ws) \end{cases} \tag{1}$$

with all the w_i's and w in $\mathbb{Z}$.

For any solution $\mathbf{e}$ of the subset problem, taking $w_i = e_i$, $w = -1$, we get $\mathcal{L}(B)$ contains a corresponding lattice vector $\mathbf{e}^{'} = (e_1^{'} \dots e_n^{'}, -\frac{1}{2}, 0)$ with $e_i^{'} = e_i - \frac{1}{2} \in \{-\frac{1}{2}, \frac{1}{2}\}$. Obviously, $\|\mathbf{e}^{'}\|_p = \frac{1}{2}(n+1)^{\frac{1}{p}}$.

On the other hand, it is easy to know that any $\mathbf{y} = (y_1, y_2 \dots y_n, y_{n+1}, y_{n+2})^T \in \mathcal{L}(B)$ of the form

$$\begin{cases} y_i \in \{-\frac{1}{2}w, \frac{1}{2}w\} & (i = 1, 2 \dots n) \\ y_{n+1} = -\frac{1}{2}w \\ y_{n+2} = 0 \end{cases}$$

where $w \in \mathbb{Z}\backslash\{0\}$ yields an solution $(y_1 - \frac{1}{2}, y_2 - \frac{1}{2}, \cdots, y_n - \frac{1}{2})$ of the RSSP. Thus, we define the solution set of the subset sum instance

$$S_n = \{ \quad w(y_1, y_2 \dots y_n, -\frac{1}{2}, 0)^T \quad | \quad |y_i| = \frac{1}{2}, \quad w \in \mathbb{Z}\backslash\{0\} \quad \}.$$

Then $\pm\mathbf{e}^{'} \in S_n$.

By querying the l_p-norm SVP oracle with $\mathcal{L}(B)$, we get a non-zero shortest vector $\mathbf{x}$. If $\mathbf{x} \in S_n$, then we can recover one solution of the RSSP. So the failure possibility is at most

$$P = \Pr(\exists \mathbf{x} \in \mathcal{L}(B) \quad \text{s.t.} \quad 0 < \|\mathbf{x}\|_p \leq \|\mathbf{e}^{'}\|_p \quad , \mathbf{x} \notin S_n).$$

For $\mathbf{x} \in \mathcal{L}(B)$ with $\|\mathbf{x}\|_p \leq \|\mathbf{e}^{'}\|_p = \frac{1}{2}(n+1)^{\frac{1}{p}}, \mathbf{x} \notin S_n$, we have $x_{n+2} = 0$ since $N > \frac{1}{2}(n+1)^{\frac{1}{p}}$, which implies

$$\sum_{i=1}^{n} w_i a_i + ws = 0. \tag{2}$$

If w is odd, then $\mathbf{x} \notin \mathbb{Z}^{n+2}$ and $|x_i| \geq \frac{1}{2}$ for $i = 1, 2 \ldots n+1$ by (1). Noticing that $\|\mathbf{x}\|_p \leq \frac{1}{2}(n+1)^{\frac{1}{p}}$, we must have $|x_i| = \frac{1}{2}$ and $w = \pm 1$, which means $\mathbf{x} \in S_n$ in this case.

Thus w is even and $\mathbf{x} \in \mathbb{Z}^{n+2}$. Using $x_i = w_i + \frac{1}{2}w$ and $x_{n+1} = \frac{1}{2}w$, together with (2), we have

$$\sum_{i=1}^{n} x_i a_i + 2x_{n+1}s - x_{n+1}\sum_{i=1}^{n} a_i = 0.$$

As a result, the above probability P can be bounded as

$$\begin{aligned}
P &= \Pr(\exists \mathbf{x} \in \mathcal{L}(B) \quad \text{s.t.} \quad 0 < \|\mathbf{x}\|_p \leq \frac{1}{2}(n+1)^{\frac{1}{p}}, \quad \mathbf{x} \notin S_n) \\
&\leq \Pr(\exists \mathbf{x} \in \mathbb{Z}^{n+1} \quad \text{s.t.} \quad 0 < \|\mathbf{x}\|_p \leq \frac{1}{2}(n+1)^{\frac{1}{p}}, \\
&\qquad \sum_{i=1}^{n} x_i a_i + 2x_{n+1}s - x_{n+1}\sum_{i=1}^{n} a_i = 0, (\mathbf{x}^T, 0)^T \notin S_n) \\
&\leq \Pr(\sum_{i=1}^{n} x_i a_i + 2x_{n+1}s - x_{n+1}\sum_{i=1}^{n} a_i = 0 : 0 < \|\mathbf{x}\|_p \leq \frac{1}{2}(n+1)^{\frac{1}{p}}, (\mathbf{x}^T, 0)^T \notin S_n) \\
&\quad \cdot \#\{\mathbf{x} \in \mathbb{Z}^{n+1} \quad | \quad \|\mathbf{x}\|_p \leq \frac{1}{2}(n+1)^{\frac{1}{p}}\}.
\end{aligned}$$

For any solution $\mathbf{e}$, we have $s = \sum_{i=1}^{n} e_i a_i$. Taking $z_i = x_i + 2x_{n+1}e_i - x_{n+1}$, we get

$$\sum_{i=1}^{n} x_i a_i + 2x_{n+1}s - x_{n+1}\sum_{i=1}^{n} a_i = 0 \Longleftrightarrow \sum_{i=1}^{n} z_i a_i = 0.$$

So we have

$$P \leq \Pr(\sum_{i=1}^{n} z_i a_i = 0, \quad (\mathbf{x}^T, 0)^T \notin S_n) \cdot \#\{\mathbf{x} \in \mathbb{Z}^{n+1} \quad | \quad \|\mathbf{x}\|_p \leq \frac{1}{2}(n+1)^{\frac{1}{p}}\}.$$

We next show that there exists a j s.t. $z_j \neq 0$. For contradiction, if all $z_j = 0$, then $x_j = (1 - 2e_j)x_{n+1}$. Hence $|x_j| = |x_{n+1}|$ since $e_j \in \{0, 1\}$. By $0 < \|\mathbf{x}\|_p \leq \frac{1}{2}(n+1)^{\frac{1}{p}}$, we know that $0 < x_j < \frac{1}{2}$, which contradicts that x_j's are integer. So there exists a j s.t. $z_j \neq 0$. Let $z^{'} = -\sum_{i \neq j} z_i a_i / z_j$, then

$$
\begin{aligned}
\Pr(\sum_{i=1}^{n} z_i a_i = 0, (\mathbf{x}^T, 0)^T \notin S_n) &= \Pr(\sum_{i=1}^{n} z_i a_i = 0, z_j \neq 0) \\
&= \Pr(a_j = z^{'}) \\
&= \sum_{k=1}^{A} \Pr(a_j = z^{'} | z^{'} = k) \cdot \Pr(z^{'} = k) \\
&= \sum_{k=1}^{A} \Pr(a_j = k) \cdot \Pr(z^{'} = k) \\
&= \frac{1}{A} \sum_{k=1}^{A} \Pr(z^{'} = k) \\
&\leq \frac{1}{A}.
\end{aligned}
$$

Now we obtain

$$P \leq \frac{1}{A} \cdot \#\{\mathbf{x} \in \mathbb{Z}^{n+1} | \|\mathbf{x}\|_p \leq \frac{1}{2}(n+1)^{\frac{1}{p}}\}. \tag{3}$$

By Theorem 1, we can bound P as

$$P \leq \frac{2^{c_p(n+1)}}{A} = \frac{2^{c_p(n+1)}}{2^{(n/d)}}$$

When $d < 1/c_p \triangleq \delta_p$, P is exponentially small on n, meaning almost all random subset sum instances with density less than δ_p can be solved by l_p-norm SVP oracle. Hence we get the following theorem.

Theorem 2. *For $p \in \mathbb{Z}^+$ and large enough n, let A be a positive integer, $a_i (1 \leq i \leq n)$ be independently uniformly random integers between 1 and A, $\boldsymbol{e} = (e_1, e_2, \cdots, e_n)$ be arbitrary non-zero vector in $\{0,1\}^n$, and $s = \sum_{i=1}^{n} a_i e_i$. If the density*

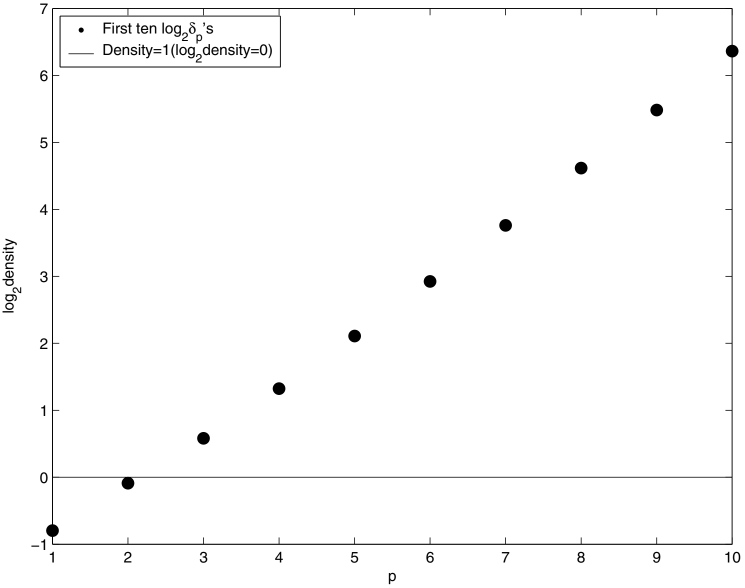
$$d = \frac{n}{\log_2 A} < \delta_p = \begin{cases} 0.5761, p = 1 \\ 0.9408, p = 2 \\ 1/(\frac{1}{2^p} \log_2(2^{p+1} - 2) + \log_2(1 + \frac{1}{(2^p - 1)(1 - (\frac{1}{2^{p+1} - 2})^{(2^p - 1)})})), p \geq 3 \end{cases} \tag{4}$$

then with probability exponentially close to 1, the subset sum problem defined by $a_i (1 \leq i \leq n)$ and s can be solved in polynomial time with a single call to an l_p-norm SVP oracle.

The table below gives the values of δ_p for p from one to ten.

p	1	2	3	4	5	6	7	8	9	10
δ_p	0.5761	0.9408	1.4957	2.5013	4.3127	7.5907	13.564	24.521	44.750	82.302

We also plot the ten $\log_2 \delta_p$'s values in the following picture.



Roughly speaking, the asymptotic form for δ_p is $2^p/(p+2)$. It's easy to see that the upper bound δ_p will go increasingly to infinity when p tends to infinity, which implies that an l_p-norm SVP oracle with larger p will help to solve more random subset sum problems. Another interesting phenomenon is that we can solve the RSSP with density one with the l_p-norm SVP oracle with $p \geq 3$ but we can not solve them with l_2-norm SVP oracle by now. It seems that the hardness of l_p-norm SVP is not decreasing as p gets larger.

4 Conclusion

In this paper, we generalize the classical probabilistic reduction from random subset sum problem to l_2-norm SVP to the case for l_p-norm. For any $p \in \mathbb{Z}^+$, we can use an l_p-norm SVP oracle to solve almost all random subset sum problem with density bounded by δ_p. Since δ_p increases as p gets bigger, an l_p-norm SVP oracle with larger p will help to solve more random subset sum problems. Moreover, an l_p-norm SVP oracle with $p \geq 3$ can help solve almost all random subset sum instances with density one, which are thought to be the most difficult instances.

Acknowledgement. We thank the anonymous referees for putting forward their excellent suggestions on how to improve the presentation of this paper.

References

1. Ajtai, M.: Gennerating hard instances of lattice problems. In: STOC 1996, pp. 99–108. ACM Press, New York (1996)
2. Ajtai, M., Dwork, C.: A public-key cryptosystem with worst-case/average-case equivalence. In: STOC 1997, pp. 284–293. ACM Press, New York (1997)

3. Ajtai, M.: The shortest vector problem in L2 is NP-hard for randomized reductions(extended abstract). In: 30th Annual ACM Symposium on Theory of Computing, pp. 266–275. ACM Press, New York (1998)
4. Babai, L.: On Lovasz' lattice reduction and the nearest lattice point problem. Combinatorica 6(1), 1–13 (1986)
5. Coster, M.J., Joux, A., Lamacchia, B.A., Odlyzko, A.M., Schnorr, C.P., Stern, J.: An improved low-density subset sum algorithm. Computational Complexity 2, 111–128 (1992)
6. Dadush, D., Peikert, C., Vempala, S.: Enumerative lattice algorithms in any norm via M -ellipsoid coverings. In: FOCS 2011, pp. 580–589. IEEE Computer Society Press (2011)
7. Frieze, A.M.: On the Lagarias-Odlyzko algorithm for the subset sum problem. SIAM J. Comput. 18, 550–558 (1989)
8. Gentry, C.: Fully homomorphic encryption using ideal lattices. In: STOC 2009, pp. 169–178. ACM Press, New York (2009)
9. Gentry, C., Peikert, C., Vaikuntanathan, V.: Trapdoors for hard lattices and new cryptographic constructions. In: STOC 2008, pp. 197–206. ACM Press, New York (2008)
10. Goldreich, D., Micciancio, D., Safra, S., Seifert, J.P.: Approximating shortest lattice vectors is not harder than approximating closest lattice vectors. Information Processing Letters 71(2), 55–61 (1999)
11. Hoffstein, J., Pipher, J., Silverman, J.H.: NTRU: A Ring-Based Public Key Cryptosystem. In: Buhler, J.P. (ed.) ANTS 1998. LNCS, vol. 1423, pp. 267–288. Springer, Heidelberg (1998)
12. Hoffstein, J., Howgrave-Graham, N., Pipher, J., Silverman, J.H., Whyte, W.: NTRUSIGN: Digital Signatures Using the NTRU Lattice. In: Joye, M. (ed.) CT-RSA 2003. LNCS, vol. 2612, pp. 122–140. Springer, Heidelberg (2003)
13. Impagliazzo, R., Naor, M.: Efficient Cryptographic Schemes Provably as Secure as Subset Sum. Journal of Cryptology 9, 199–216 (1996)
14. Lenstra, A.K., Lenstra Jr., H.W., Lovasz, L.: Factoring polynomials with rational coefficients. Mathematische Annalen 261, 513–534 (1982)
15. Lagarias, J.C., Odlyzko, A.M.: Solving low-density subset sum problems. J. Assoc. Comp. Mach. 32(1), 229–246 (1985)
16. Lyubashevsky, V., Micciancio, D., Peikert, C., Rosen, A.: SWIFFT: A modest proposal for FFT hashing. In: Nyberg, K. (ed.) FSE 2008. LNCS, vol. 5086, pp. 54–72. Springer, Heidelberg (2008)
17. Micciancio, D., Goldwasser, S.: Complexity of Lattice Problems: A Cryptography Perspective. Kluwer Academic Publishes (2002)
18. Micciancio, D., Voulgaris, P.: A deterministic single exponential time algorithm for most lattice problems based on Voronoi cell computations. In: STOC 2010, pp. 351–358. ACM Press, New York (2010)
19. Micciancio, D.: Inapproximability of the Shortest Vector Problem: Toward a Deterministic Reduction. Theory of Computing 8(1), 487–512 (2012)
20. Regev, O.: Lattices in computer science. Lecture notes of a course given in Tel Aviv University (2004)
21. Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. In: STOC 2005, pp. 84–93. ACM Press, New York (2005)
22. Regev, O., Rosen, R.: Lattice problems and norm embeddings. In: STOC 2006, pp. 447–456. ACM Press, New York (2006)