JACOBI SYMBOL PARITY CHECKING ALGORITHM FOR SUBSET PRODUCT

TREY LI

ABSTRACT. It is well-known that the subset product problem is NP-hard. We give a probabilistic polynomial time algorithm for the special case of high $\mathbb{F}_2$-rank.

1. INTRODUCTION

The subset product problem (SP) [GJ79] is about solving a multivariable exponential equation

$$\prod_{i=1}^{n} a_i^{x_i} = X$$

for a binary solution $(x_1,\ldots,x_n) \in \{0,1\}^n$, where $a_1,\ldots,a_n,X \in \mathbb{Z}$. Andrew Yao shown its NP-completeness in a private communication in 1978 [GJ79, p. 224, p. 325]. This means that worst-case SP cannot be solved in polynomial time unless P = NP. We show that the special case of SP with *characteristic matrix* of $\mathbb{F}_2$-rank $\geq n - \log_2(n^c)$ for some constant c can be solved in probabilistic polynomial time.

2. CHARACTERISTIC MATRIX

Let $p_1,\ldots,p_m$ be the prime factors of $a_1,\ldots,a_n$ in ascending order. We call a matrix $A \in \mathbb{Z}^{m\times n}$ the *characteristic matrix* of the SP if

$$a_i = \prod_{j=1}^{m} p_j^{A_{j,i}}$$

for all $i \in [n]$. In other words, $a_1,\ldots,a_n$ are products of primes selected by the columns of A from $p_1,\ldots,p_m$. Also notice that m is possibly greater than, equal to, or smaller than n.

We call the row rank (over any possible field) of the characteristic matrix the *rank* (over the same field) of the SP. The rank (over a specified field) is an invariant of an SP instance.

3. ALGORITHM

Step 1. Choose $k \geq m$ random integers $s_1,\ldots,s_k$. Reduce the equation

$$\prod_{i=1}^{n} a_i^{x_i} = X$$

to k modular equations of the form

$$\prod_{i=1}^{n} a_i^{x_i} \equiv X \pmod{s_j},$$

for $j \in [k]$.

This is the 2[nd] paper of the series. Previously: [Li22].

Date: October 2, 2022.

Email: treyquantum@gmail.com

Step 2. Take the Jacobi symbols of $a_1,\dots,a_n,X$ for each equation to get k equations

$$\prod_{i=1}^{n}\left(\frac{a_i}{s_j}\right)^{x_i}=\left(\frac{X}{s_j}\right),$$

for $j\in[k]$.

Step 3. Extract from the above system a matrix equation

$$Bx\equiv b \pmod 2,$$

where $B\in\{0,1\}^{k\times n}$, $B_{j,i}=[1-(a_i/s_j)]/2$; and $b\in\{0,1\}^k$, $b_j=[1-(X/s_j)]/2$. In other words, the entries of B and b are obtained by mapping the Jacobi symbols from -1 to 1 and 1 to 0.

We call B the characteristic matrix of the Jacobi symbol matrix $\{(a_i/s_j)\}_{j\in[k],i\in[n]}$; and b the characteristic vector of the Jacobi symbol vector $((X/s_j))_{j\in[k]}$.

Step 4. Search in the solution set of $Bx\equiv b \pmod 2$ for one that satisfies $\prod_{i=1}^{n}a_i^{x_i}=X$.

4. MAXIMIZING RANK(B)

Note that all solutions to the SP are solutions to the equation

$$Bx\equiv b \pmod 2.$$

We want the rank of B over $\mathbb{F}_2$ to be as high as possible to reduce the searching complexity of Step 4.

Let $P\in\{0,1\}^{k\times m}$ be the characteristic matrix of the Jacobi symbol matrix $\{(p_i/s_j)\}_{j\in[k],i\in[m]}$ with respect to the primes p_i. We have

$$B=PA,$$

where $A=\{A_{j,i}\}_{m\times n}$ is the characteristic matrix of the SP. The rank of B over $\mathbb{F}_2$ is

$$\operatorname{rank}_{\mathbb{F}_2}(B)=\operatorname{rank}_{\mathbb{F}_2}(PA)\le\operatorname{rank}_{\mathbb{F}_2}(A)\le\min\{m,n\},$$

where $\operatorname{rank}_{\mathbb{F}_2}(B)$ achieves its maximum value $\operatorname{rank}_{\mathbb{F}_2}(A)$ when P achieves its maximum rank m.

We show the existence of m integers $s_1,\dots,s_m$ such that the characteristic matrix $P\in\{0,1\}^{m\times m}$ of the Jacobi symbol matrix $\{(p_i/s_j)\}_{j\in[m],i\in[m]}$ is of full $\mathbb{F}_2$-rank. It is sufficient to prove the following lemma, which is about achieving an arbitrary row of an arbitrary $P\in\{0,1\}^{m\times m}$.

LEMMA 1. Let $p_1,\dots,p_m$ be distinct primes. For any vector $v\in\{-1,1\}^m$, there exists an integer s such that the vector of Jacobi symbols $((p_1/s),\dots,(p_m/s))=v$.

Proof. Case (1). All p_i are odd. By the law of quadratic reciprocity, the Jacobi symbols satisfy

$$\left(\frac{p_i}{s}\right)=\left(\frac{s}{p_i}\right)$$

if and only if $p_i\equiv 1 \pmod 4$ or $s\equiv 1 \pmod 4$. Take $s\equiv 1 \pmod 4$. Then

$$\left(\left(\frac{p_1}{s}\right),\dots,\left(\frac{p_m}{s}\right)\right)=v$$

if s satisfies the following $m+1$ equations:

$$s\equiv 1 \pmod 4;\ \text{ and }\left(\frac{s}{p_i}\right)=v_i,\ \text{for } i\in[m].$$

Since p_i are odd primes, the Jacobi symbols (s/p_i) are Legendre symbols. So if $v_i = 1$ then we can define the corresponding equation to be

$$s \equiv 1 \pmod{p_i},$$

because 1 is always a quadratic residue. Otherwise if $v_i = -1$ then we define the corresponding equation to be

$$s \equiv r_i \pmod{p_i},$$

where r_i is any quadratic non-residue modulo p_i[1]. So the $m+1$ equations boil down to

$$s \equiv 1 \pmod{4}; \text{ and}$$
$$s \equiv 1 \pmod{p_i} \text{ if } v_i = 1, \text{ or } s \equiv r_i \pmod{p_i} \text{ if } v_i = -1, \text{ for } i \in [m].$$

By the Chinese remainder theorem (CRT), there is a unique solution $s \in \mathbb{Z}_{4\prod_{i=1}^{m} p_i}$.

Case (2). There is an even prime $p_a = 2$ and $v_a = 1$, for some $a \in [m]$. Note that the Jacobi symbol

$$\left(\frac{2}{s}\right) = 1$$

if $s \equiv 1, 7 \pmod{8}$. We take $s \equiv 1 \pmod{8}$. This also implies that $s \equiv 1 \pmod{4}$ and thus

$$\left(\frac{p_i}{s}\right) = \left(\frac{s}{p_i}\right)$$

for the odd primes p_i. Then the m equations that s needs to satisfy is

$$s \equiv 1 \pmod{8}; \text{ and } \left(\frac{s}{p_i}\right) = v_i, \text{ for } i \in [m], i \neq a.$$

They boil down to

$$s \equiv 1 \pmod{8}; \text{ and}$$
$$s \equiv 1 \pmod{p_i} \text{ if } v_i = 1, \text{ or}$$
$$s \equiv r_i \pmod{p_i} \text{ if } v_i = -1, \text{ for } i \in [m], i \neq a.$$

By CRT there is a unique solution $s \in \mathbb{Z}_{8\prod_{i\in[m], i\neq a} p_i}$.

Case (3). There is an even prime $p_a = 2$ and $v_a = -1$, for some $a \in [m]$. Note that the Jacobi symbol

$$\left(\frac{2}{s}\right) = -1$$

if $s \equiv 3, 5 \pmod{8}$. We take $s \equiv 3 \pmod{8}$. This also implies $s \equiv 3 \pmod{4}$.

Again, notice that for the odd primes p_i, if $p_i \equiv 1 \pmod{4}$, then

$$\left(\frac{p_i}{s}\right) = \left(\frac{s}{p_i}\right);$$

hence for $\left(\frac{p_i}{s}\right) = v_i$ it suffices to require

$$\left(\frac{s}{p_i}\right) = v_i.$$

[1]If necessary, it is easy to find r_i by sampling random elements from $\mathbb{Z}_{p_i}^{\times}$ because half of the elements in $\mathbb{Z}_{p_i}^{\times}$ are quadratic non-residues.

Otherwise if $p_i \equiv s \equiv 3 \pmod 4$, then

$$\left(\frac{p_i}{s}\right) = -\left(\frac{s}{p_i}\right);$$

hence for $\left(\frac{p_i}{s}\right) = v_i$ it suffices to require

$$\left(\frac{s}{p_i}\right) = -v_i.$$

Hence the m equations s needs to satisfy is

$$s \equiv 3 \pmod 8; \text{ and}$$
$$\left(\frac{s}{p_i}\right) = v_i \text{ if } p_i \equiv 1 \pmod 4, \text{ or}$$
$$\left(\frac{s}{p_i}\right) = -v_i \text{ if } p_i \equiv 3 \pmod 4, \text{ for } i \in [m], i \neq a.$$

They boil down to

$$\begin{aligned} &s \equiv 3 \pmod 8; \text{ and}\\ s \equiv 1 \pmod{p_i} \text{ if: } &[v_i = 1 \text{ and } p_i \equiv 1 \pmod 4] \text{ or}\\ &[v_i = -1 \text{ and } p_i \equiv 3 \pmod 4], \text{ or}\\ s \equiv r_i \pmod{p_i} \text{ if: } &[v_i = -1 \text{ and } p_i \equiv 1 \pmod 4] \text{ or}\\ &[v_i = 1 \text{ and } p_i \equiv 3 \pmod 4], \text{ for } i \in [m], i \neq a. \end{aligned}$$

By CRT there is a unique solution $s \in \mathbb{Z}_{8\prod_{i\in[m],i\neq a} p_i}$. □

However, since we do not have the prime factors $p_1,\dots,p_m$ of the integers $a_1,\dots,a_n$, we cannot find s deterministically as in the proof of Lemma 1. We therefore choose $k \geq m$ random integers $s_1,\dots,s_k$ as in Step 1, and expect that for a polynomial size k the matrix $P \in \{0,1\}^{k\times m}$ achieves its full rank m. Then B achieves its maximum rank $\text{rank}_{\mathbb{F}_2}(A)$.

5. Theorem

We state the theorem in terms of average-case SP with uniform characteristic matrix. The conclusion about best-case SP with high rank characteristic matrix, as stated in Abstract and Introduction, is implied.

Theorem 1. Let $m,n,d \in \mathbb{N}$ with $m \geq n$, and $d \geq 2$ even. There exists a probabilistic polynomial time algorithm that solves SP with uniform characteristic matrix $A \in \mathbb{Z}_d^{m\times n}$ with respect to random prime factors $p_1,\dots,p_m$ with probability $\gtrapprox \prod_{i=m-n+1}^{m}(1-1/2^i)$.

Proof. Consider the algorithm given by Section 3. It is clear that the time complexity is polynomially in n assuming polynomial many solutions to $Bx \equiv b \pmod 2$. Now we prove that this happens with probability $\gtrapprox \prod_{i=m-n+1}^{m}(1-1/2^i)$.

By the randomness of $p_1,\dots,p_m$ we expect that by polynomially many random integers $s_1,\dots,s_k$, the $\mathbb{F}_2$-rank of $P \in \{0,1\}^{k\times m}$ achieves m with overwhelming probability. I.e., $\text{rank}_{\mathbb{F}_2}(B) = \text{rank}_{\mathbb{F}_2}(A)$ with overwhelming probability.

Again, the probability [Lan93; Ber80; BS06] that a uniform matrix in $\mathbb{F}_2^{m\times n}$ with $m \geq n$ is of full $\mathbb{F}_2$-rank is

$$p = \prod_{i=m-n+1}^{m} \left(1 - \frac{1}{2^i}\right).$$

Now $A \in \mathbb{Z}_d^{m\times n}$ and d is even. I.e., the entries of A are from $\{0,\ldots,d-1\}$, where half numbers are odd and half numbers are even. Hence $A \pmod 2$ is uniform over $\mathbb{F}_2$ and that it is of full $\mathbb{F}_2$-rank with probability p.

If A is really full rank, we solve for the unique x and check if it gives a solution to the SP. Else if A is not full rank but close to full rank, namely $2^{n-\mathrm{rank}_{\mathbb{F}_2}(A)} \leq n^c$ for some constant c, we can still check all solutions of $Bx \equiv b \pmod 2$ and see if there is one that satisfies the SP. Hence the probability of solving SP is $\geq p$ assuming $\mathrm{rank}_{\mathbb{F}_2}(B) = \mathrm{rank}_{\mathbb{F}_2}(A)$.

Combining the overwhelming probability of $\mathrm{rank}_{\mathbb{F}_2}(B) = \mathrm{rank}_{\mathbb{F}_2}(A)$, we have the claimed probability of $\gtrapprox p$. □

The following Corollary gives a better idea of what Theorem 1 means.

Corollary 1. Let $m, n, d \in \mathbb{N}$ with $m \geq 2n$, and $d \geq 2$ even. There exists a probabilistic polynomial time algorithm that solves average-case SP with uniform characteristic matrix $A \in {}_d^{m\times n}$ with respect to random prime factors $p_1, \ldots, p_m$ with overwhelming probability.

Proof. Simply plug $m \geq 2n$ in p we have that $\mathrm{rank}_{\mathbb{F}_2}(A) = n$ with probability

$$p = \prod_{i=m-n+1}^{m} \left(1 - \frac{1}{2^i}\right) \geq \prod_{i=n+1}^{2n} \left(1 - \frac{1}{2^i}\right) > \left(1 - \frac{1}{2^n}\right)^n,$$

which is overwhelming in n. □

References

[Ber80] E.R. Berlekamp. "The technology of error-correcting codes". In: *Proceedings of the IEEE* 68.5 (1980), pp. 564–593. DOI: 10.1109/PROC.1980.11696.

[BS06] Ian F Blake and Chris Studholme. "Properties of random matrices and applications". In: *Unpublished report available at http://www. cs. toronto. edu/˜cvs/coding* (2006).

[GJ79] Michael R Garey and David S Johnson. *Computers and intractability*. Vol. 174. freeman San Francisco, 1979.

[Lan93] Georg Landsberg. "Ueber eine Anzahlbestimmung und eine damit zusammenhängende Reihe." In: *Journal für die reine und angewandte Mathematik* 111 (1893), pp. 87–88. URL: http://eudml.org/doc/148874.

[Li22] Trey Li. "Subset Product with Errors over Unique Factorization Domains and Ideal Class Groups of Dedekind Domains". 1st paper of the series. 2022, October 1.