低ノイズLPN問題におけるBKWアルゴリズムの解析

上中谷　健†　　國廣　昇‡　　高安　敦††

東京大学
277-8561 千葉県柏市柏の葉 5-1-5
†ken_kaminakaya@mist.i.u-tokyo.ac.jp
‡kunihiro@k.u-tokyo.ac.jp
††a-takayasu@it.k.u-tokyo.ac.jp

あらまし　LPN 問題は量子計算機に耐性を持つと期待され，低ノイズの LPN 仮定は公開鍵暗号の構成など盛んに研究されている．この LPN 問題を BKW アルゴリズムは初めて準指数時間で解いたが，準指数個と多くのサンプルを必要とする．Lyubashevsky はハッシュ関数を用いることでサンプル数を多項式個に減らしたが，計算時間が遅くなってしまった．本研究は既存研究のサンプル数および計算時間をより厳密に評価することで，両者の関係を式で表し，両者の関係を明確化した．さらに，計算時間にノイズを考慮することで，低ノイズの場合では，計算時間を同じオーダーのままサンプル数を多項式オーダーにできることを示した．

Analysis of BKW Algorithm for Low Noise LPN Problem

Ken Kaminakaya†　　Noboru Kunihiro‡　　Atsushi Takayasu††

†The University of Tokyo.
5-1-5 Kashiwanoha, Kashiwa, Chiba 2778561, JAPAN
†ken_kaminakaya@mist.i.u-tokyo.ac.jp
‡kunihiro@k.u-tokyo.ac.jp
††a-takayasu@it.k.u-tokyo.ac.jp

Abstract　Recently, LPN problem has been studied actively and public key encryption based on LPN assumption was proposed. BKW algorithm can solve LPN problem in subexponential time, but the number of required samples is subexponential. Lyubashevsky proposed sample amplification which makes the number of samples polynomial, though the running time is longer. In this paper, we express the equation of the relation between them as we estimate them more tightly. In addition, we show that Lyubashevsky ' s method can solve low noise LPN problem requiring polynomial number of samples as the running time is the same as that of BKW algorithm.

1　はじめに

1.1　背景

LPN 問題 (Learning Parity with Noise problem) とは Blum 等 [2] によって提唱された問題である．LPN 問題とは，Q 個のサンプル $(\boldsymbol{g}_i, \boldsymbol{s}\cdot\boldsymbol{g}_i+e_i(\mathrm{mod}2))$ から $\boldsymbol{s}$ を復元する問題である．ただし，$\boldsymbol{s}, \boldsymbol{g}_i \in \{0,1\}^n$ はランダムに生成されたベクトル，e_i は確率 p のベルヌーイ分布に従って生成されたノイズである．量子計算機上で素因数分解などが多項式時間で解けるようになる

中，LPN 問題は，多項式時間で解くアルゴリズムが存在していないため，量子計算機でも計算量的に困難な問題であると期待されている．そのため，LPN 問題の困難性を利用した暗号が盛んに研究されている．また，$n = 512, p = \frac{1}{8}$ のとき 80 ビットセキュリティであるとされていたが，それを否定する結果が現れ，困難性は不明瞭のままである．以上より，暗号の安全性に使用されている LPN 問題の計算量の理論的限界を評価することは重要である．

さらに，低ノイズ下でも LPN 問題は計算量的に困難であるという仮定を利用した研究が近頃盛んに行われている．暗号として利用する場合，誤り率が大きいと復号が困難であるため，低ノイズの仮定が必要なのである．たとえば，Kiltz 等 [8] は低ノイズ下の LPN 仮定を基にした IND-CCA2 安全な公開鍵暗号を提案し，Döttling[4] は低ノイズの LPN 仮定を基にした KDM-CPA 安全な公開鍵暗号を提案した．両者の論文で，低ノイズとは $p = O(1/\sqrt{n})$ である．よって，これらの安全性を評価するため，ノイズの大きさを考慮した評価が必要である．

LPN 問題を総当たり法よりも速く解くアルゴリズムは，BKW アルゴリズム，Arora-Ge アルゴリズム，シンドローム復号問題として解くアルゴリズムの 3 種類が存在する．本論文は，Blum 等 [1] が提案した初めての準指数時間アルゴリズムであり，数多く研究されている BKW アルゴリズムを対象とする．BKW アルゴリズムの計算時間は $2^{O(n/\log n)}$ と比較的短いが，解くために必要なサンプル数は $2^{O(n/\log n)}$ と多くなってしまった．そこで，Lyubashevsky[10] は，少ないサンプルを線形結合することで BKW アルゴリズムに必要な数だけ増やすことを提案した．それにより $n^{1+\epsilon}$ 個のサンプルで解くことができるが，計算時間は $2^{O(n/\log\log n)}$ と遅くなった．ただし，ϵ は正の定数である．上中谷と國廣 [12] は Lyubashevsky の手法にパラメータ w を導入した．w は，何個のサンプルを線形結合するかを表す値であり，サンプル数 Q と計算時間 T に影響を与える．[12] は w を動かすことで Blum 等 [1] と Lyubashevsky[10] の結果の間となる (Q, T) を見つけ，それをつなげることでトレードオフの関係があると推測した．BKW アルゴリズムを改良した手法はほかにも Bogos 等 [3]，Fossorier 等 [5]，Guo 等 [6]，Levieil と Fouque[9] があるが，オーダーは BKW アルゴリズムと変わらない．

1.2 研究成果

本研究は，拡張した Lyubashevsky の手法 [12] におけるサンプル数 Q と計算時間 T の関係を明確化した．まず，[12] ではパラメータ w を導入した際の T が厳密に評価されていなかったのに対し，本研究は T の具体的な値を求めた．次に，Q と T の連立方程式から w を消去することで，定数 C_1, C_2 を用いて，

$$Q = C_1(\log T - C_2)2^{2n/C_1(\log T - C_2)}$$

という Q と T の関係式を求めた．最後にパラメータを動かし，Q と T の関係をグラフ化した．このグラフにより，[12] のトレードオフの関係であるという推測を否定することができた．

さらに，既存研究では $p = O(1)$ として計算時間の評価にノイズの大きさを考慮していないのに対して，本研究はノイズの大きさを考慮して評価を行った．その結果，LPN 問題が低ノイズ，すなわち，$p = 1/n^{\Omega(1)}$ であるとき，計算時間 T を BKW アルゴリズムと同じ $2^{O(n/\log n)}$ のままで，サンプル数 Q を多項式オーダーにできることを示した．

1.3 構成

前節の研究成果を示すにあたり，まず，必要な表記と概念を 2 章で定義する．3 章では，本論文が解析した既存手法を紹介する．具体的には，3.1 節では根幹である BKW アルゴリズムを，3.2 節ではサンプル数を減らす Lyubashevsky の手法を，3.3 節では拡張した Lyubashevsky の手法を紹介する．前節のサンプル数と計算時間の関係を 4.1 節で明確化し，LPN 問題が低ノイズの場合を 4.2 節で評価し，研究成果を示す．最後に，5 章で本論文をまとめる．

2 準備

2.1 表記

本論文で使用する表記を本節で定義する:

- $\boldsymbol{s} \leftarrow \{0,1\}^n$: $\boldsymbol{s}$ はランダムで選ばれた長さ n のバイナリベクトル
- Ber_p : パラメータ p のベルヌーイ分布 $(0 < p < 1/2)$
- $e \leftarrow \mathrm{Ber}_p$: $\Pr[e = 1] = p$ または $\Pr[e = 0] = 1 - p$
- $\boldsymbol{s} \cdot \boldsymbol{x}$: mod 2 上のベクトル $\boldsymbol{s}$ と $\boldsymbol{x}$ の内積
- $\boldsymbol{u}_i$: 成分 i が 1 の単位ベクトル
- $\oplus$: 排他的論理和（xor）
- $\boldsymbol{x}_1 \oplus \boldsymbol{x}_2$: ビット毎の $\boldsymbol{x}_1$ と $\boldsymbol{x}_2$ の xor
- $\mathrm{Hw}(\boldsymbol{x})$: ベクトル $\boldsymbol{x}$ のハミング重み
- log : 底が 2 の対数関数
- ln : 自然対数.

2.2 Learning Parity with Noise

LPN 問題 (Learning Parity with Noise problem) を定義するため，以下のサンプル $(\boldsymbol{g}, z)$ を出力するオラクル $\mathcal{O}_{\boldsymbol{s},p}$ を定義する：

$$\{(\boldsymbol{g}, z) = (\boldsymbol{g}, \boldsymbol{s} \cdot \boldsymbol{g} \oplus e) | \boldsymbol{g} \leftarrow \{0,1\}^n, e \leftarrow \mathrm{Ber}_p\}.$$

ただし，n は入力 $\boldsymbol{s}$ の長さとした．このオラクルを用いてアルゴリズム $\mathcal{A}$ が LPN 問題を解くということを以下で定義する.

定義 1. セキュリティパラメータ n を入力として，アルゴリズム $\mathcal{A}$ が時間 T，メモリ M，オラクル $\mathcal{O}_{\boldsymbol{s},p}$ から得た Q 個のサンプルでベクトル $\boldsymbol{s}$ の b ビットを予測できる確率が θ 以上のとき，すなわち，

$$\Pr[\mathcal{A}^{\mathcal{O}_{\boldsymbol{s},p}}(1^n) = (s_1, \cdots, s_b) | \boldsymbol{s} \leftarrow \{0,1\}^n] \geq \theta$$

のとき，アルゴリズム $\mathcal{A}$ は (Q, T, M, θ, b) で $\mathrm{LPN}_{n,p}$ 問題を解くという.

オラクル $\mathcal{O}_{\boldsymbol{s},p}$ に Q 回アクセスしたとき，Q 個のサンプル $(g_1, z_1), \cdots, (g_Q, z_Q)$ を得ることができる．ただし，$z_i = \boldsymbol{s} \cdot \boldsymbol{g_i} \oplus e_i$ が成り立っている．この $z_i, e_i (i = 1, \cdots, Q)$ をベクトル $\boldsymbol{z}, \boldsymbol{e}$ として，$\boldsymbol{g_i} (i = 1, \cdots, Q)$ を $n \times Q$ の行列 G として考えることで，

$$\boldsymbol{z} = G^T \boldsymbol{s} + \boldsymbol{e} \pmod 2 \qquad (1)$$

が成立する．つまり，$\mathrm{LPN}_{n,p}$ 問題とはランダムに選ばれたバイナリ行列 G と，Ber_p に従うノイズ付線型方程式 (1) によって得られる $\boldsymbol{z}$ から，$\boldsymbol{s}$ を復元するという問題である.

2.3 統計的距離

本論文は Lyubashevsky[10] と同様に統計的距離を用いて証明する．そのため，統計的距離の定義と命題を本節で紹介する．なお，命題の証明については [11] を参照されたい.

定義 2. X, Y を可算集合 S 上の確率変数とする．このとき，X, Y の統計的距離を $\Delta[X; Y]$ と表記し，以下のように定義する：

$$\Delta[X; Y] = \frac{1}{2} \sum_{s \in S} |\Pr[X = s] - \Pr[Y = s]|.$$

命題 3. $X_1, \cdots, X_k$ および $Y_1, \cdots, Y_k$ を独立な確率変数列とする．このとき，以下が成立する:

$$\Delta[(X_1, \cdots, X_k); (Y_1, \cdots, Y_k)] \leq \sum_{i=1}^{k} \Delta[X_i; Y_i].$$

命題 4. X, Y を可算集合 S 上の確率変数とする．任意の述語 $f : S \to \{0,1\}$ において，以下が成立する：

$$|\Pr[f(X) = 1] - \Pr[f(Y) = 1]| \leq \Delta[X; Y].$$

3 既存研究

本論文の解析に必要な既存研究を本章で紹介する.

3.1 BKW アルゴリズム [1]

Blum 等 [1] が提案した BKW アルゴリズムは主に 2 つの段階で構成されている．第 1 段階では，オラクルから得たベクトルに対してブロッ

Algorithm 1 BKW アルゴリズム [1]

入力：$ab \geq n$ を満たす自然数 a, b
　　N', L：自然数
repeat
　L 個のサンプル $(\boldsymbol{g}_j, z_j)$ をオラクル $\mathcal{O}_{\boldsymbol{s},p}$ から得る (その集合を S とする)
　$\boldsymbol{g}_j$ を b ビット毎に a 個のブロックに分割
　for $i = 0$ to $a-2$ **do**
　　集合 S をベクトルの第 $a-i$ ブロックの値毎に $S_1, \cdots, S_{2^b}$ に分割する
　　for $l = 1$ to 2^b **do**
　　S_l からランダムに $(\boldsymbol{g}_*, z_*)$ を選ぶ
　　S_l のすべての $(\boldsymbol{g}, z)$ に対して，
　　$(\boldsymbol{g}, z) \leftarrow (\boldsymbol{g} \oplus \boldsymbol{g}_*, z \oplus z_*)$
　　$(\boldsymbol{g}_*, z_*)$ を S_l から削除
　　end for
　　$S \leftarrow S_1 \cup \cdots \cup S_{2^b}$
　end for
until $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_b$ を含むサンプルがそれぞれ N' 個出力される
for $i = 1$ to b **do**
　N' 個のサンプル $(\boldsymbol{u}_i, z)$ のうち，z で多数決を取りその結果を s_i として出力
end for

クごとにガウスの消去法を行い，単位ベクトルを十分な個数生成する．第 2 段階では，生成された単位ベクトルを含むサンプルを基に多数決を行い，1 ビットの予測を行う．これらの動作を表したものが Algorithm 1 である．また，このアルゴリズムでは下記の定理が成り立つ．

定理 5 ([1])**.** 自然数 a, b は $ab \geq n$ を満たす．$\mathrm{LPN}_{n,p}$ 問題を解くために必要な計算時間およびサンプル数は $\mathrm{poly}\left(\left(\frac{1}{1-2p}\right)^{2^a}, 2^b\right)$ である．

この定理に対して，$a = \delta \log n$ $(0 < \delta < 1)$, $b = \lceil \frac{n}{a} \rceil$, $p = O(1)$ のとき，$Q = 2^{O(n/\log n)}$, $T = 2^{O(n/\log n)}$ となる．

3.2 Lyubashevsky の手法 [10]

Lyubashevsky[10] は BKW アルゴリズムの必要なサンプル数を減らすために Impagliazzo と Zuckerman[7] が示した補題 6 を使用した．

補題 6 ([7])**.** 集合 X, Y は $X \subseteq \{0,1\}^Q$，$|X| > 2^{2n}$，$Y = \{0,1\}^n$ を満たす．ただし，Q は正の定数．$G \in \{0,1\}^{n \times Q}$ の i 番目の列ベクトルを $\boldsymbol{g}_i$ とする．さらに，X から Y へのハッシュ関数 $h_G(\boldsymbol{x}) = \bigoplus_{i=1}^{Q} x_i \boldsymbol{g}_i$ とその集合族 $\mathcal{H} = \{h_G | G \in \{0,1\}^{n \times Q}\}$ を定義する．このとき，h_G が $\mathcal{H}$ からランダムに選ばれた関数ならば，少なくとも $1 - 2^{-\frac{n}{4}}$ の確率で $\Delta[h_G(\boldsymbol{x}); U] \leq 2^{-\frac{n}{4}}$ が成立する．ただし，$\boldsymbol{x}$ は X からランダムに選ばれたベクトルで，U は Y 上の一様分布とする．

正の定数 ϵ に対し，$Q = n^{1+\epsilon}, X = \{\boldsymbol{x} \in \{0,1\}^Q | \mathrm{Hw}(\boldsymbol{x}) = \frac{2n}{\epsilon \log n}\}$ とする．また，集合 X からランダムに選んだ $\boldsymbol{x}$ を入力とする関数 $h_G(\boldsymbol{x}), h_{\boldsymbol{z}}(\boldsymbol{x})$ を $h_G(\boldsymbol{x}) = \bigoplus_i x_i \boldsymbol{g}_i$，$h_{\boldsymbol{z}}(\boldsymbol{x}) = \bigoplus_i x_i z_i$ と定義する．[10] は，オラクル $\mathcal{O}_{\boldsymbol{s},p}$ から得た Q 個のサンプル $(\boldsymbol{g}_i, z_i)$ と $\boldsymbol{x}_j \leftarrow X$ を入力として，N 個のサンプル

$$(\boldsymbol{g}'_j, z'_j) = (h_G(\boldsymbol{x}_j), h_{\boldsymbol{z}}(\boldsymbol{x}_j)) \quad (j = 1, \cdots, N) \tag{2}$$

を生成し，BKW アルゴリズムをブラックボックスで用いることで，LPN 問題を解くことを提案した．この手法により，$n^{1+\epsilon}$ 個のサンプルを最大 $|X|$ 個のサンプルに増やすことができる．また，$|X| > 2^{2n}$ かつ，$G = [\boldsymbol{g}_1 \cdots \boldsymbol{g}_Q]$ がオラクルによりランダムに生成されているので，補題 6 を適用することができ，さらに命題 3 と命題 4 を用いることで適用後の成功確率もほぼ変わらないことが示せる．計算時間は次節で示す．

3.3 Lyubashevsky 方式の拡張 [12]

上中谷と國廣 [12] は $w < Q$ を導入して，

$$X = \{\boldsymbol{x} \in \{0,1\}^Q \mid \mathrm{Hw}(\boldsymbol{x}) = w\}$$

として Lyubashevsky の手法を評価した．ただし，パラメータ w は自然数である．このとき，

$$|X| = \binom{Q}{w} > \left(\frac{Q}{w}\right)^w$$

であるから，

$$Q \geq w 2^{\frac{2n}{w}} \Longrightarrow |X| > 2^{2n}$$

となる．すなわち，U を $\{0,1\}^n$ 上の一様分布としたとき，$Q \geq w2^{\frac{2n}{w}}$ ならば，補題 6 より，$\Delta[h_G(\boldsymbol{x}); U] \leq 2^{-\frac{n}{4}}$ が成立する．

計算時間の評価にあたり，[10] は下記の補題 7 と balls-and-bins 理論を用いて解析したが，[12] では補題 7 だけを用いてより厳密に評価した．

補題 7 (Piling-Up Lemma)．オラクル $\mathcal{O}_{\boldsymbol{s},p}$ に l 回アクセスして，l 個のサンプル $(\boldsymbol{g}_1, z_1), \cdots, (\boldsymbol{g}_l, z_l)$ を得られたとする．このとき，

$$\Pr[z_1 \oplus \cdots \oplus z_l = \boldsymbol{s} \cdot (\boldsymbol{g}_1 \oplus \cdots \oplus \boldsymbol{g}_l)] = \frac{1}{2} + \frac{1}{2}(1-2p)^l$$

が成立する．

式 (2) のサンプルを入力とした BKW アルゴリズムの多数決で使用されるサンプル $(\boldsymbol{u}_k, z_*)$ は 2^{a-1} 個の $(\boldsymbol{g}'_j, z'_j)$ の排他的論理和を取ったものであるため，$\mathrm{Hw}(\boldsymbol{y}) = 2^{a-1}$ を満たすベクトル $\boldsymbol{y} = \{y_j\}$ を用いて，

$$\boldsymbol{u}_k = \bigoplus_{j=1}^{N} y_j \boldsymbol{g}'_j, \quad z_* = \bigoplus_{j=1}^{N} y_j z'_j$$

となる．$G' = [\boldsymbol{g}'_1 \cdots \boldsymbol{g}'_N]$, $\boldsymbol{z}' = \{z'_j\}$ とすれば，

$$\boldsymbol{u}_k = G'\boldsymbol{y} \pmod 2, \quad z_* = \boldsymbol{z}' \cdot \boldsymbol{y} \tag{3}$$

が成立する．$M_x = [\boldsymbol{x}_1 \cdots \boldsymbol{x}_N]$ とすれば，式 (2) は

$$G' = GM_x, \quad \boldsymbol{z}' = M_x^T \boldsymbol{z} \pmod 2$$

となる．これを式 (3) に代入することで，

$$\boldsymbol{u}_k = GM_x\boldsymbol{y}, \quad z_* = \boldsymbol{z}^T M_x \boldsymbol{y} \pmod 2 \tag{4}$$

が成立する．$\boldsymbol{\alpha} = M_x\boldsymbol{y}$ とすれば，式 (4) は

$$\boldsymbol{u}_k = \bigoplus_{i=1}^{Q} \alpha_i \boldsymbol{g}_i, \quad z_* = \bigoplus_{i=1}^{Q} \alpha_i z_i$$

となる．つまり，$(\boldsymbol{u}_k, z_*)$ は，オラクル $\mathcal{O}_{\boldsymbol{s},p}$ より与えられた Q 個のサンプルのうち，$\mathrm{Hw}(\boldsymbol{\alpha})$ 個のサンプルの排他的論理和を取ったものである．よって，補題 7 より，

$$\begin{aligned}\Pr[s_k = z_*] &= \Pr[\boldsymbol{s} \cdot \boldsymbol{u}_k = z_*] \\ &= \frac{1}{2} + \frac{1}{2}(1-2p)^{\mathrm{Hw}(\boldsymbol{\alpha})}\end{aligned}$$

となる．$\mathrm{Hw}(\boldsymbol{x}_j) = w$ かつ $\mathrm{Hw}(\boldsymbol{y}) = 2^{a-1}$ より，

$$\mathrm{Hw}(\boldsymbol{\alpha}) \leq w2^{a-1}$$

であるから，

$$\Pr[s_k = z_*] \geq \frac{1}{2} + \frac{1}{2}(1-2p)^{w2^{a-1}} \tag{5}$$

が成立する．これは，1 ビット s_k を z_* と予測したときの正解確率の下限を示している．

BKW アルゴリズムでは，式 (5) の成功確率を上げるために，N' 個あるサンプル $(\boldsymbol{u}_k, z_*)$ で多数決を取る．その評価ために確率変数 $Y_1, \cdots, Y_{N'}$ を以下のように定義する：

$$Y_i = \begin{cases} 1 & s_k\text{の予測が不正解} \\ 0 & s_k\text{の予測が正解} \end{cases} \quad (i = 1, \cdots, N').$$

式 (5) で $\epsilon' \geq (1-2p)^{w2^{a-1}}/2$ としたとき，

$$\Pr[Y_i = 0] = \frac{1}{2} + \epsilon' \quad (i = 1, \cdots, N')$$

が成立する．このとき，$S_{N'} = \sum_{i=1}^{N'} Y_i$ は N' 回の予測を行ったときの不正解数を表す．ベルヌーイ分布の Hoeffding の不等式 (15) より，

$$\begin{aligned}\Pr\left[S_{N'} \geq \frac{N'}{2}\right] &= \Pr\left[S_{N'} \geq N'\left(\frac{1}{2} - \epsilon' + \epsilon'\right)\right] \\ &\leq \exp(-2N'\epsilon'^2) \\ &\leq \exp\left(-\frac{1}{2}N'(1-2p)^{w2^a}\right)\end{aligned} \tag{6}$$

となる．Hoeffding の不等式については付録 A を参照されたい．よって，$N' = \mathrm{poly}((\frac{1}{1-2p})^{w2^a}, b)$ であれば，多数決で不正解する確率は漸近的に 0 に近づく．

以上より，$a2^b$ 個のサンプルに排他的論理和を取ることを N' 回行うため，

$$T = \mathrm{poly}\left(\left(\frac{1}{1-2p}\right)^{w2^a}, 2^b\right), \tag{7}$$

$$Q = w2^{\frac{2n}{w}} \tag{8}$$

となる．また，Lyubashevsky の手法 [10] は上式の $w = \frac{2n}{\epsilon \log n}, a = \delta \log\log n, b = \lceil \frac{n}{a} \rceil$ のときであるため，$T = 2^{O(n/\log\log n)}$ となる．

4　サンプル数と計算時間

4.1　サンプル数と計算時間の関係

前節で示したように拡張されたLyubashevskyの手法[12]では計算時間Tにpoly関数が使用されているため，関係を立式できなかった．そこで，本節ではTの具体的な値を求め，[12]におけるサンプル数と計算時間の関係を明確化する．

$L = a2^b$個のベクトルが与えられたとき，Algorithm 1の4∼12行目の作業を行った結果，単位ベクトル$\boldsymbol{u}_k$が存在すれば1を，存在しなければ0を出力する述語をfとする．[1],[12]より，

$$\Pr[f(\boldsymbol{g}_1, \cdots, \boldsymbol{g}_L) = 1] \geq 1 - \frac{1}{e}$$

が成立する．ただし，$\boldsymbol{g}_i \leftarrow \{0,1\}^n \ (i = 1, \cdots, L)$．一方，式(2)を満たす$L$個のサンプル$(\boldsymbol{g}'_j, z'_j)$において，命題3，命題4より，

$$\begin{aligned}
&|\Pr[f(\boldsymbol{g}'_1, \cdots, \boldsymbol{g}'_L) = 1] - \Pr[f(\boldsymbol{g}_1, \cdots, \boldsymbol{g}_L) = 1] \\
&\quad \leq \Delta[(\boldsymbol{g}'_1, \cdots, \boldsymbol{g}'_L); (\boldsymbol{g}_1, \cdots, \boldsymbol{g}_L)] \\
&\quad \leq \sum_{i=1}^{L} \Delta[\boldsymbol{g}'_i; \boldsymbol{g}_i] \\
&\quad \leq a2^{b-\frac{n}{4}}
\end{aligned}$$

が成立する．よって，$b < \frac{n}{4}$ならば，BKWアルゴリズムの第1段階で$(\boldsymbol{u}_k, z_*)$が生成される確率は漸近的に$1 - \frac{1}{e}$に近づく．ゆえに，第2段階においてN'個のサンプル$(\boldsymbol{u}_k, z_*)$で多数決をするためには，期待値として合計

$$N = a2^b \cdot \frac{1}{1 - \frac{1}{e}} \cdot N'$$

個のサンプル$(\boldsymbol{g}'_j, z'_j)$が必要である．

次に，N'を求める．N'が十分大きいとき，bビットの予測が正解する確率θは，式(6)より，

$$\begin{aligned}
\theta &\geq \left(1 - \exp\left(-\frac{1}{2}N'(1-2p)^{w2^a}\right)\right)^b \\
&\approx 1 - b\exp\left(-\frac{1}{2}N'(1-2p)^{w2^a}\right)
\end{aligned}$$

となる．よって，

$$N' = 2\left(\frac{1}{1-2p}\right)^{w2^a} \ln\left(\frac{b}{1-\theta}\right).$$

以上より，N個のサンプル$(\boldsymbol{g}'_j, z'_j)$に対して，$n+1$ビットの排他的論理和を$(a-1)$回行っているため，計算時間は以下のようになる：

$$T = O(naN),$$

$$N = \frac{a2^{b+1}}{1 - \frac{1}{e}} \left(\frac{1}{1-2p}\right)^{w2^a} \ln\left(\frac{b}{1-\theta}\right).$$

オーダーを外すため定数Cを導入すると，

$$T = \frac{Cna^2 2^{b+1}}{1 - \frac{1}{e}} \left(\frac{1}{1-2p}\right)^{w2^a} \ln\left(\frac{b}{1-\theta}\right)$$

となる．対数を取り，整理すると，

$$w = C_1(\log T - C_2) \tag{9}$$

となる．ただし，

$$C_1 = \frac{1}{2^a \log(\frac{1}{1-2p})},$$

$$C_2 = \log\left(\frac{Cn^2 a 2^{b+1}}{1 - \frac{1}{e}}\right) + \log\ln\left(\frac{b}{1-\theta}\right)$$

である．一方，サンプル数は式(8)であるため，式(9)を代入すると，

$$Q = C_1(\log T - C_2)2^{2n/C_1(\log T - C_2)} \tag{10}$$

となる．解析のために対数を取ると，

$$\log Q = \log(C_1(\log T - C_2)) + \frac{2n}{C_1(\log T - C_2)} \tag{11}$$

となる．式(11)では，Tが比較的小さいとき，第2項が支配項になるため$\log Q$と$\log T$は反比例の形になる一方，Tが比較的大きいとき，第1項が支配項になるためQは$\log T$に比例することがわかる．これにより，[12]のトレードオフの関係になるという推測は否定された．

式(10)において$n = 1024, p = \frac{1}{8}, \theta = \frac{1}{2}, b = \lceil \frac{n}{a} \rceil, C = 1$のとき，図1のようになる．図1には，式(7)の2つの項がほぼ等しくなる値である

$$a = \log\left(\frac{n}{p'w}\right) - \log\log\left(\frac{n}{p'w}\right) \tag{12}$$

がグラフ化されていて，最適値に近い値となっている．これより式(12)は準最適値となることがわかる．ただし，$p' = -\log(1-2p)$とした．

式(10)の関係はBKWアルゴリズムを改良した研究にも適用できると考えられる．たとえば，Bogos等[3]について付録Bに載せておいた．

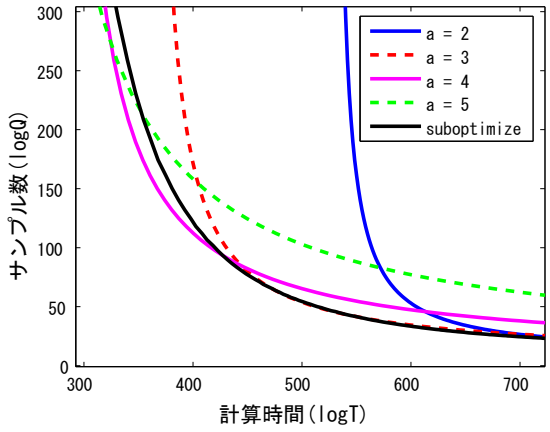

図 1: LPN 問題 ($n = 1024, p = \frac{1}{8}$) を解くために必要なサンプル数 Q と計算時間 T の関係

4.2 低ノイズの場合

既存研究では $p = O(1)$ として T を評価したものが多数である．しかし，定理 5 からも明らかなように，LPN 問題のノイズが小さくなればなるほど，計算時間は短くなる．また，実際 LPN 仮定を使用するにあたり，低ノイズ，すなわち $p = O(1/\sqrt{n})$ とした研究も行われている ([4],[8])．このことから，p を考慮した計算時間 T を評価する必要がある．

p が小さいとき，$1 - 2p \approx \exp(-2p)$ であるため，$p' \approx 2p$ である. よって，式 (12) を代入した結果以下のようになる：

$$T = 2^{O(n/(\log \frac{n}{pw} - \log\log \frac{n}{pw}))}.$$

上式に $w = n/\log n$ を代入すると，

$$Q = n^3/\log n,$$

$$T = 2^{O(n/(\log\log n - \log p))} \qquad (13)$$

となる．

$p = O(1)$ または $p = 1/(\log n)^{\Omega(1)}$ ならば，式 (13) の指数の分母の支配項は $\log\log n$ になるため，

$$T = 2^{O(n/\log\log n)}$$

となる．これは Lyubashevsky[10] の Q は多項式オーダー，$T = 2^{O(n/\log\log n)}$ と同じ結果となる．

一方，低ノイズ，すなわち $p = 1/n^{\Omega(1)}$ の場合を考える．定数 c を用いて，$p \leq 1/n^c$ であるため，$-\log p \geq c\log n$ が成立する．よって，低ノイズの場合の式 (13) の指数の分母の支配項は $c\log n$ である．よって，

$$T = 2^{O(n/\log n)}$$

となる．すなわち，Lyubashevsky の手法を低ノイズの場合に適用すると，計算時間は BKW アルゴリズムと同じオーダーのままサンプル数を多項式オーダーにできることが示された．

以上の議論は w が $n/\log n$ の定数倍であっても成立する．

5 まとめ

本研究は拡張された Lyubashevsky の手法のサンプル数と計算時間に着目した．既存研究では，計算時間が明確な式で表されてなかったため，サンプル数と計算時間の関係はトレードオフの関係になるという推測までしかできていなかったが，本研究は計算時間を立式することで，サンプル数と計算時間の関係を具体的な式で表すことに成功した．その式を解析およびグラフ化することで，サンプル数はある値から増加することがわかり，トレードオフの関係を否定することができた．さらに，ノイズを変化させたとき，サンプル数が多項式オーダーのときの計算時間が減っていくことに着目し，低ノイズの場合を解析した．その結果，あるパラメータでは，BKW アルゴリズムと同じ $2^{O(n/\log n)}$ のままでサンプル数を多項式オーダーにできることを示した．

参考文献

[1] Avrim Blum, Adam Kalai, and Hal Wasserman, “Noise-tolerant learning, the parity problem, and the statistical query model,” J.ACM 50(4):506-519, 2003.

[2] Avrim Blum, Merrick L. Furst, Michael J. Kearns, and Richard J. Lipton, “Crypto-

graphic primitives based on hard learning problems," CRYPTO 1993: 278-291.

[3] Sonia Bogos, Florian Tramèr, and Serge Vaudenay, "On solving LPN using BKW and variants," Cryptology ePrint Archive: Report 2015/049.

[4] Nico Döttling, "Low noise LPN: KDM secure public key encryption and sample amplification," PKC 2015:604-626.

[5] Marc P.C. Fossorier, Miodrag J. Mihaljević, Hideki Imai, Yang Cui and Kanta Matsuura, "A novel algorithm for solving the LPN problem and its application to security evaluation of the HB protocol for RFID authentication," IACR, 2006.

[6] Qian Guo, Thomas Johansson, and Carl Löndahl, "Solving LPN using covering codes," Asiacrypt 2014: 1-20.

[7] Russell Impagliazzo and David Zuckerman, "How to recycle random bits," IEEE, 1989.

[8] Eike Kiltz, Daniel Masny and Krzysztof Pietrzak, "Simple chosen-ciphertext security from low-noise LPN," PKC 2014:1-18.

[9] Éric Levieil and Pierre-Alain Fouque, "An improved LPN algorithm," SCN 2006: 348-359.

[10] Vadim Lyubashevsky, "The parity problem in the presence of noise, decoding random linear codes, and the subset sum problem," RANDOM 2005.

[11] Victor Shoup, "A computational introduction to number theory and algebra second edition," Cambridge university press, 2009.

[12] 上中谷健，國廣昇，"LPN 問題に対する BKW アルゴリズムの拡張," SCIS 2015.

A Hoeffding の不等式

$X_1, \cdots, X_n$ は独立な確率変数，かつ X_i の値域は有限区間 $[a_i, b_i]$ とする $(i = 1, \cdots, n)$．さらに，$S_n = \sum_{i=1}^{n} X_i$ とする．このとき，任意の $\epsilon > 0$ に対し，

$$\Pr[S_n - \mathrm{E}[S_n] \geq n\epsilon] \leq \exp\left(-\frac{2\epsilon^2 n^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right) \tag{14}$$

が成立する．

特に，$X_1, \cdots, X_n \in [0, 1]$ が Ber_p に従うとき，式 (14) は

$$\Pr[S_n \geq n(p + \epsilon)] \leq \exp(-2\epsilon^2 n) \tag{15}$$

となる．

B Bogos 等 [3] への適用

Bogos 等 [3] は BKW アルゴリズムの多数決に必要なサンプルを最初にすべて用意しておくことで，削除するサンプルを最小限にし，以下の定理を証明した．

定理 8 ([3])．改良した BKW アルゴリズムは $(Q = 2^{b+1}(1-2p)^{-2^a}\log(\frac{b}{1-\theta}) + (a-1)2^b, T = O(naQ), M = nQ, \theta, b)$ で $\mathrm{LPN}_{n,p}$ 問題を解く．ただし，a, b は $ab \geq n$ を満たす自然数である．

このアルゴリズムに拡張した Lyubashevsky の手法 [12] を適用したとき，計算時間は

$$T' = C' n a 2^{b+1} \left(\frac{1}{1-2p}\right)^{w2^a} \ln\left(\frac{b}{1-\theta}\right)$$

となる．上式と式 (8) の連立方程式より，

$$Q' = C'_1(\log T' - C'_2)2^{2n/C'_1(\log T' - C'_2)}$$

となる．ただし，

$$C'_1 = \frac{1}{2^a \log(\frac{1}{1-2p})},$$

$$C'_2 = \log(C' n a 2^{b+1}) + \log\ln\left(\frac{b}{1-\theta}\right)$$

である．