

Revised Method for Sampling Coefficient Vector of GNR-enumeration Solution

Gholam Reza Moghissi

Department of ICT, Malek-Ashtar University of Technology, Tehran, Iran
E-mail: fumoghissi@chmail.ir

Ali Payandeh

Department of ICT, Malek-Ashtar University of Technology, Tehran, Iran

Received: 02 October 2021; Revised: 18 November 2021; Accepted: 25 December 2021; Published: 08 August 2022

Abstract: For selecting security parameters in lattice-based cryptographic primitives, the exact manner of BKZ algorithm (as total cost and specification of output basis) should be estimated in high block sizes. The simulations of BKZ are used to predict (estimate) the exact manner of BKZ algorithm which cannot be studied by practical running of BKZ algorithm for higher block sizes. Sampling method of GNR (Gamma-Nguyen-Regev) enumeration solution vector $\mathbf{v}$ is one of the main components of designing BKZ-simulation and it includes two phases: sampling the norm of solution vector $\mathbf{v}$ and sampling corresponding coefficient vectors. Our work, by Moghissi and Payandeh in 2021, entitled as “*Better Sampling Method of Enumeration Solution for BKZ-Simulation*”, introduces a simple and efficient idea for sampling the norm and coefficient vectors of GNR enumeration solution $\mathbf{v}$. This paper proposes much better analysis for approximating the expected value and variance of the entries of these coefficient vectors. By this analysis, our previous idea for sampling the coefficient vectors is revised, which means that the expected value and variance of every entry in these coefficient vectors sampled by our new sampling method, are more close to the expected value and variance of corresponding entries in original sampling method, while these new sampled coefficient vectors include no violation from main condition of GNR bounding function (i.e., our new sampling method is not a rejection sampling).

Index Terms: BKZ Simulation, GNR Enumeration, Coefficient Vector, Sampling Method, Expected Value, Variance.

1. Introduction

Lattice reduction is the main phase of most lattice attacks. BKZ reduction is the current main practical lattice reduction algorithm [1,2]. For determining the security parameters in lattice cryptographic primitives (such as the work in [3]), the total cost and output quality of BKZ algorithm should be estimated in high block sizes. For predicting the manner of BKZ in higher block sizes, practical running is not the way, therefore BKZ-simulations are introduced (e.g., the simulation by Shi Bai et al. [4] is one of the last versions of BKZ-simulation). Designing a BKZ-simulation with GNR-pruned enumeration includes some necessary components such as enumeration radius, generation of bounding function, estimation of success probability, LLL simulation, estimation of GNR enumeration cost, sampling method for enumeration solution and simulation of updating GSO [5]. However other SVP-solvers (such as the works in [6-8]), instead of GNR-enumeration, can be considered in designing BKZ-simulations, (to the best of our knowledge) previous researches in design of BKZ-simulation just focus on BKZ with enumeration function. Also, design of pure version of BKZ-simulation (with GNR-enumeration) can be extended simply into the simulation of different versions of BKZ (such as BKZ with progressive block size [9] or progressive success probability [10] or pressed BKZ [4]).

One of the main components in design of BKZ-simulation with GNR enumeration is sampling the enumeration solution $\mathbf{v}$. Enumeration solution $\mathbf{v}$ in former BKZ-simulations is often considered in the form of the expected value of the norm of solution $\|\mathbf{v}\|$. GNR enumeration function in BKZ algorithm over a lattice block $\mathcal{L}_{[b_1, \dots, b_k]}$ returns the coefficient vector $\mathbf{y}$, which can be used to compute the solution vector $\mathbf{v}$ by linear combination of vectors of this lattice block as $\mathbf{v} = \sum_{i=1}^k y_i \mathbf{b}_i$ [5]. Our work in [5] tries to introduce an approximate sampling method for enumeration solution $\mathbf{v}$ which samples both norm and coefficient vectors of enumeration solution. Our idea in Lemma 3 from [5] introduces a suitable method for sampling the norm of enumeration solution $\mathbf{v}$. In other side, to the best of our knowledge, first precise and explicit analysis of sampling the coefficient vector $\mathbf{y}$ is introduced in [5] and following sampling methods of coefficient vectors are proposed in [5]:

- Our sampling method 1 (which is introduced in Claim 1 and Lemma 11 from [5]; see the pseudo-code in Algorithm 5 from [5]) is defined just for piecewise-linear bounding functions with condition of small enumeration radius or small (static) success probability. In other words, the condition that the expected number of solutions in GNR enumeration (i.e., dynamic success frequency) belongs to $\approx O(1)$; However our results in [5] verify our sampling method 1, this version of sampling is not too accurate for much small success probability;
- Our sampling method 2 (see the pseudo-code in Algorithm 6 from [5]) is defined just for piecewise-linear bounding functions with any enumeration radius and any success probability; However our results in [5] verify our sampling method 2, this version of sampling is not too accurate for enumeration radius of $R \approx \|v\|$ and much small success probability;
- Our sampling method 3 (see the pseudo-code in Algorithm 6 from [5]) is adapted for any bounding function with any enumeration radius and any success probability (i.e., with no constraint); Our test results in [5] show that the probability distribution (and expected value) of $X_i = \sum_{l=1}^i z_l^2 / R^2$ for different types of bounding functions with same success probability P is not sufficiently similar, unless for big dynamic success frequency (i.e., $f_{succ} \gg 1$). Accordingly, this version of sampling is not much accurate.

As mentioned, our idea in Lemma 3 from [5] is accepted in this paper, but our sampling methods of 1, 2 and 3 cannot be accepted as exact sampling methods of coefficient vectors and would be revised in this paper.

The accuracy of our sampling methods of 1, 2 and 3 in [5] is just determined based on the expected value of each entry in coefficient vector of enumeration solution! As shown in test results of this paper and [5], when the success probability of bounding function is much small or enumeration radius tends to the solution norm (i.e., $R \approx \|v\|$), the proposed sampling method 1, 2 and 3 would be less accurate. In this paper, a novel idea for approximating this expected value is introduced which is more close to the expected value of corresponding entries in basic rejection sampling in Algorithm 6 from [2]. In other side, there is no analysis on variance of each entry in coefficient vector of enumeration solution in [5], while our test results in this paper show disappointing variance for these three sampling methods! Also a novel idea for approximating this variance is introduced which is more close to corresponding entries in basic rejection sampling in Algorithm 6 from [2]. Moreover there is no analysis on the number of violations from main condition of bounding function at each entry in coefficient vector of enumeration solution in [5], while our test results in this paper show surprising much number of violations for these three sampling methods! Finally in this paper, our sampling method 1, 2 and 3 (for sampling coefficient vectors) are revised, in the way that, our new method is not a rejection sampling (i.e., it includes no violation of main condition of GNR bounding function at each entry of coefficient vectors), also the expected value and variance of each entry of coefficient vectors tend to our new exact approximation of expected value and variance of corresponding entries in Algorithm 6 from [2]. This is worthy of noting that for designing a BKZ-simulation, our novel idea for approximating the expected value of corresponding entries in Algorithm 6 from [2] is sufficiently exact (instead of “Our Sampling Method 5”), while its variance and number of violations at each entry are absolutely zero. In fact, using our new sampling method just tries to inject more reasonable randomness to the BKZ-simulation.

The remainder of this paper is organized as follows. Section 2 is dedicated to essential background for understanding our contributions in this paper. Our contributions in design of sampling method of coefficient vectors of GNR enumeration solution would be introduced in section 3. Our test results for verifying our claims and value of our contributions in this paper are introduced in section 4. Finally, in section 5, the conclusions and further studies for this work are expressed.

2. Background

In this section, a sufficient background is introduced to make this work easy to be understood. At first of this section, some basic preliminaries for this paper are introduced. Next, some necessary information on GNR-enumeration and coefficient vector of GNR-enumeration solution is introduced.

2.1. Basic Preliminaries

A lattice in this paper is defined as follows:

Lattices. A lattice is a set of points in the n -dimensional space [11], in the way that for given n -linearly independent vectors $b_1, \dots, b_n \in \mathbb{R}^m$, the lattice can be generated by them as the following set of vectors:

$$\mathcal{L}(b_1, \dots, b_n) = \{\sum_{i=1}^n x_i b_i : x_i \in \mathbb{Z}\}. \quad (1)$$

The set of vectors of $[b_1, \dots, b_n]$ is known as the lattice basis shown by column matrix of B . Also a lattice block of $[b_j, b_{j+1}, \dots, b_k]$ is shown by $B_{[j,k]}$ or $\mathcal{L}_{[b_j, \dots, b_k]}$ or $\mathcal{L}_{[j,k]}$. For cryptographic applications, this is assumed to $b_i \in \mathbb{Z}^m$. The rank and dimension of lattice $\mathcal{L}(B)$ are shown respectively by n and m .

Euclidean norm. The length or norm of a lattice vector $v = (v_1, \dots, v_m)$ is measured by $\|v\| = \sqrt[2]{v_1^2 + \dots + v_m^2}$.
Fundamental Domain. For a lattice $\mathcal{L}(B)$, the fundamental domain is defined as following set:

$$\mathcal{F}(\mathcal{L}) = \{t_1 b_1 + t_2 b_2 + \dots + t_n b_n : 0 \leq t_i < 1\}. \quad (2)$$

Volume of Lattice. The volume of a lattice $\mathcal{L}(B)$ is defined by the volume of the parallelepiped of fundamental domain $\mathcal{F}(\mathcal{L})$ which is computed as follows:

$$\text{Vol}(\mathcal{L}(B)) = \text{Vol}(\mathcal{F}(\mathcal{L}(B))) = |\det B|. \quad (3)$$

One of the main heuristic in lattice theory is Gaussian Heuristic, which estimates the number of points in a set S .

Heuristic 1 (Gaussian Heuristic). “Given a lattice $\mathcal{L}$ and a set S , the number of points in $S \cap \mathcal{L}$ is approximated by $\text{Vol}(S)/\text{Vol}(\mathcal{L})$ ” [1];

By using Gaussian Heuristic, if lattice $\mathcal{L}$ is limited in a central ball with radius of $R = \lambda_1(\mathcal{L})$, then it is expected that there is at least one lattice vector in $\text{Ball}_n(R)$ with radius R , which is the shortest vector. Therefore, the value of $\lambda_1(\mathcal{L})$ can be estimated by Gaussian Heuristic of lattice $GH(\mathcal{L})$ as follows (by using sterling approximation):

$$GH(\mathcal{L}) = \left(\frac{\text{Vol}(\mathcal{L}(B))}{\text{Vol}(\text{Ball}_n(1))} \right)^{\frac{1}{n}} \approx \sqrt{\frac{n}{2\pi e}} (\det B)^{\frac{1}{n}}. \quad (4)$$

Also, Gram-Schmidt Orthogonalization and projected sub-lattices are other fundamental definitions in BKZ reduction.

Orthogonal projection ($\pi_i(\dots)$). For a given lattice basis $B = (b_1, b_2, \dots, b_n)$, the orthogonal projection $\pi_i(\dots)$ is defined as follows:

$$\pi_i: \mathbb{R}^m \mapsto \text{span}(b_1, \dots, b_{i-1})^\perp, \quad \text{where } 1 \leq i \leq n.$$

Gram-Schmidt Orthogonal basis (GSO basis). For given lattice basis $B = (b_1, b_2, \dots, b_n)$, the Gram-Schmidt Orthogonal basis $B^* = (b_1^*, b_2^*, \dots, b_n^*)$ is defined as follows:

$$\begin{aligned} \pi_i(b_i) &= b_i^* = b_i - \sum_{j=1}^{i-1} \mu_{i,j} b_j^*, \\ \text{where } \mu_{i,j} &= \frac{b_i \cdot b_j^*}{\|b_j^*\|^2} \text{ and } 1 \leq j < i \leq n. \end{aligned} \quad (5)$$

Note: The parameter of b_i^* refers to i -th vector of GSO basis of B^* .

Also in our statistical analysis, Gamma distribution is massively used:

Gamma distribution. The Gamma distribution is a two-parameter and continuous probability distribution which is defined as follows (for input shape-parameter of k and scale-parameter of θ):

$$\text{Gamma}(x; k, \theta) = \frac{x^{k-1} e^{-x/\theta}}{\Gamma(k) \theta^k}, \quad \text{where } x > 0. \quad (6)$$

The mean and variance in Gamma distribution respectively are determined by $k\theta$ and $k\theta^2$.

Note: In this paper, the expected value and variance of random variable X respectively are shown by $E[X]$ and $\mathcal{V}[X]$ (i.e., for random variable of x , the notation of $\mathcal{V}[x]$ and the notation of $\sigma^2[x]$ are equivalent). Also, the standard deviation is shown by notation of $\text{STD}[X] \approx \sqrt{\mathcal{V}[X]}$.

2.2. GNR-Enumeration function

There are many hard problems in lattice theory, where the shortest vector problem (SVP) is the basic one. For given lattice basis B , solvers of SVP try to find the shortest nonzero vector in this lattice. One of the main SVP-solvers is enumeration function which searches the solution vector in a tree. Different pruning method is introduced for enumeration function.

Full-enumeration. For initial radius R , the tree of full-enumeration enumerates all lattice points in n -dimensional ball of radius R (i.e., no pruning is used in this type of enumeration function).

The sound pruning technique (also named as GNR-pruning), which is introduced by Gama, Nguyen and Regev, uses the concept of cylinder-intersection. The cylinder-intersection, bounding function and GNR-pruning are defined as follows:

Cylinder-intersection. The l -dimensional cylinder-intersection with radius $(R_1, \dots, R_l)$ is defined as following set

[1]:

$$C_{R_1, \dots, R_l} = \{(x_1, \dots, x_l) \in \mathbb{R}^l, \forall 1 \leq i \leq l, \sum_{t=1}^i x_t^2 \leq R_i^2\}. \quad (7)$$

Bounding function. The vector of $\mathcal{R}$: $0 \leq \mathcal{R}_1 \leq \mathcal{R}_2 \leq \dots \leq \mathcal{R}_\beta = 1$, when multiplied by initial radius R , defines a bounded cylinder-intersections with radius $(R_1, \dots, R_l) = (R \cdot \mathcal{R}_1, \dots, R \cdot \mathcal{R}_l)$ for $1 \leq l \leq \beta$, and consequently is used to prune the enumeration trees [1].

Note: In this paper, the entry i in bounding function $\mathcal{R}$ is shown by $\mathcal{R}[i]$ and $\mathcal{R}_i$.

GNR-pruning (Sound pruning). For a lattice block of $B_{[j,k]} = [b_j, b_{j+1}, \dots, b_k]$ and the coefficient vector $x \in \mathbb{Z}^\beta$, this pruning replaces the inequalities of $\|\pi_{k+1-i}(x \cdot B_{[j,k]})\| \leq R$ for $1 \leq i \leq k - j + 1$ as a bounded ball (in full-enumeration) by $\|\pi_{k+1-i}(x \cdot B_{[j,k]})\| \leq \mathcal{R}_i R$, where $0 \leq \mathcal{R}_1 \leq \dots \leq \mathcal{R}_{k-j+1} = 1$ as a cylinder-intersection [2].

The pseudo-code of the sound pruned enumeration (GNR-enumeration) can be studied in Appendix B from paper [1]. Success probability is one of the main features of bounding function which is defined as follows [1]:

Success probability of bounding function. For any lattice block of $[b_j, b_{j+1}, \dots, b_k]$, initial enumeration radius R and bounding function $\mathcal{R}$, if there is just one lattice vector v in n -dimensional ball with radius of R (i.e., $\|v\| \leq R$), the probability of finding solution vector v after GNR-pruning by $\mathcal{R}$ in enumeration tree is defined as the success probability of $\mathcal{R}$, and is shown by $p_{succ}(\mathcal{R})$.

For analysis of the success probability of GNR bounding function, Gama et al. use following heuristic assumptions (in addition to Gaussian Heuristic) [1]:

Heuristic 2. “The distribution of the coordinates of the target vector v , when written in the normalized Gram-Schmidt basis $(b_1^*/\|b_1^*\|, \dots, b_n^*/\|b_n^*\|)$ of the input basis, look like those of a uniformly distributed vector of norm $\|v\|$ ”.

The success probability of the bounding function $\mathcal{R}$ can be estimated by Monte Carlo simulation (see Algorithm 8 in paper [2]), however Algorithm 7 from paper [1] can be used as more efficient estimation of success probability.

At the end of this subsection, two families of bounding functions including “Step bounding function” and “Piecewise-linear bounding function” are defined as follows [1]:

Step bounding function. The Step bounding function is defined as $\mathcal{R}_i^2 = a$, for $i = 1, \dots, \beta/2$ and $\mathcal{R}_i^2 = 1$ in otherwise, where $0 < a \leq 1$ [1].

Piecewise linear pruning. The piecewise linear bounding function is defined as $\mathcal{R}_i^2 = \frac{2ia}{\beta}$, for $i = 1, \dots, \beta/2$ and $\mathcal{R}_i^2 = \left(2a - 1 + \frac{2i(1-a)}{\beta}\right)$ in otherwise, where $0 < a \leq 1$ [1].

We show in [12], if success probability of optimal bounding function is closed to %100, then the cost of GNR-enumeration pruned by this bounding function can be approximated as $\frac{1}{2^{\beta/4}}$ times the cost of full-enumeration (in consistent with the analysis in [1]). Also, we show in [12], for practical block size of $100 \leq \beta \leq 250$, the upper-bound for speedup of GNR-enumeration pruned by a piecewise-linear bounding function with much small success probability (as extreme pruning) when the lattice block is preprocessed by BKZ with enumeration would be estimated from $2^{\frac{\beta}{6.6}}$ to $2^{\frac{\beta}{4.4}}$ (and tries to reject the claimed speedup of $2^{\frac{\beta}{2}}$ by extreme-pruning over full-enumeration in [1]). Furthermore, a global search strategy, a practical search algorithm and an efficient approximate technique to generate the optimal bounding function are introduced in [14].

2.3. Coefficient Vector of GNR-Enumeration Solution

GNR enumeration function over a lattice block $\mathcal{L}_{[b_j, \dots, b_k]}$ returns the coefficient vector y , which can be used to compute the solution vector v by linear combination of vectors of this lattice block as $v = \sum_{l=j}^k y_l b_l$. The coefficient of orthonormal basis vector $z = (z_1, z_2, \dots, z_{k-j+1=d})$ in Heuristic 2 which corresponds with the target lattice vector of v can be formulated as follows (note that, $b_i^*/\|b_i^*\|$ is the i th vector of the orthonormal basis of $b_1^*/\|b_1^*\|, \dots, b_n^*/\|b_n^*\|$) [1]:

$$v = (z_1, \dots, z_d) \cdot \begin{bmatrix} b_k^*/\|b_k^*\| \\ \vdots \\ b_j^*/\|b_j^*\| \end{bmatrix} = (v_1, \dots, v_m). \quad (8)$$

The coordinates of the coefficient vector z are reversed (i.e., z_i corresponds with $b_{k-i+1}^*/\|b_{k-i+1}^*\|$), and it is clear that $\|z\| = \|v\|$ [1]. Also, the vector $u = (u_1, u_2, \dots, u_{k-j+1=d}) = (z_1/R, z_2/R, \dots, z_d/R)$ is chosen to be uniformly distributed from d -dimensional ball of the radius 1 (by the notation of $u \sim \text{Ball}_d$). By using these formulations, success probability of a GNR bounding function $\mathcal{R}$ can be formally defined as follows [1]:

$$p_{succ}(\mathcal{R}) = \Pr_{u \sim \text{Ball}_d} \left(\forall i \in [1, d], \sum_{l=1}^i u_l^2 \leq \frac{R_i^2}{R_d^2} \right) = \Pr_{u \sim \text{Ball}_d} \left(\forall i \in [1, d], \sum_{l=1}^i u_l^2 \leq \mathcal{R}_i^2 \right). \quad (9)$$

Note: Since in last block of BKZ, the size of blocks becomes less than initial block size of β , so the variable of $d = k - j + 1$ is used to emphasize this fact.

The main condition (or constraint) which defines the success probability of finding true coefficient vector of enumeration solution for an input bounding function is defined as follows:

$$(\forall i \in [1, d], \sum_{l=1}^i z_l^2 \leq \mathcal{R}_i^2 \cdot R_d^2) \equiv (\forall i \in [1, d], \sum_{l=1}^i u_l^2 \leq \frac{R_i^2}{R_d^2} = \mathcal{R}_i^2) \equiv (\forall i \in [1, d], \sum_{l=1}^i \frac{\omega_{d-l+1}}{\sum_{t=1}^d \omega_t} \leq \mathcal{R}_i^2), \quad (10)$$

where $\omega_i \leftarrow \text{Gamma}(1/2, 2)$.

The solution vector v can be written by the coefficient vector $w = \left(\frac{z_d}{\|b_1^*\|}, \dots, \frac{z_2}{\|b_{d-1}^*\|}, \frac{z_1}{\|b_d^*\|} \right)$ on GSO block as follows [5]:

$$v = (v_1, \dots, v_m) = (w_1, \dots, w_d) \cdot \begin{bmatrix} b_1^* \\ \vdots \\ b_d^* \end{bmatrix}. \quad (11)$$

Note: The solution vector v from enumeration over lattice block $\mathcal{L}_{[j,k]}$ is a GSO projected vector which is orthogonal over the previous basis vectors in $\mathcal{L}_{[1,j-1]}$ (remember that, here the notation of $\mathcal{L}_{[1,d]}$ represents $\mathcal{L}_{[j,k]}$).

By inserting the solution vector v at first of the lattice block $\mathcal{L}_{[1,d]}$ (which results in the block of $(v, b_1^*, \dots, b_d^*)$ with $d + 1$ vectors), one of the vectors from the GSO block $[v, b_1^*, \dots, b_d^*]$ should be eliminated after updating GSO coefficients of these $d + 1$ vectors. Lattice enumeration uses the integer coefficients y_i for enumerating over the projected lattice block $\mathcal{L}_{[1,d]}$, therefore the coefficients w_i in vector w depend on integer entries in the vector of y , as follows [5]:

Note: Here, the projection notation of $\pi_1(\dots)$ represents $\pi_j(\dots)$, also the dimension of b_i^* and v is m which differs from the rank of lattice block (i.e., block size of d).

$$v = (y_1, y_2, \dots, y_d) \cdot \begin{bmatrix} \pi_1(b_1) \\ \vdots \\ \pi_1(b_d) \end{bmatrix} = (y_1, y_2, \dots, y_d) \cdot \begin{bmatrix} b_1^* \\ \vdots \\ b_d^* + \sum_{i=1}^{d-1} \mu_{d,i} \cdot b_i^* \end{bmatrix}. \quad (12)$$

$$\|v\|^2 = \underbrace{(y_1 + \sum_{i=2}^d y_i \mu_{i,1})^2}_{z_d^2} \|b_1^*\|^2 + \dots + \underbrace{(y_g + \sum_{i=g+1}^d y_i \mu_{i,g})^2}_{z_{d-g+1}^2} \|b_g^*\|^2 + \dots + \underbrace{y_d^2}_{z_1^2} \|b_d^*\|^2 \Rightarrow \quad (13)$$

$$w = \left[\underbrace{y_1 + \sum_{i=2}^d y_i \mu_{i,1}}_{w_1}, \dots, \underbrace{y_g + \sum_{i=g+1}^d y_i \mu_{i,g}}_{w_g}, \dots, \underbrace{y_d}_{w_d} \right]. \quad (14)$$

Based on (14), a zero coefficient of y_i does not always result in $w_i = 0$, except for indices after the last non-zero coefficient of y_g [5]. Also by using Theorem 2 in [5], following relations can be expanded:

$$w_i = \frac{z_{d-i+1}}{\|b_i^*\|} \Rightarrow w_g = \frac{z_{d-g+1}}{\|b_g^*\|} = 1 \Rightarrow z_{d-g+1} = \|b_g^*\|. \quad (15)$$

$$w_l = \frac{z_{d-l+1}}{\|b_l^*\|} = 0 \Rightarrow z_{d-l+1} = 0, \text{ for } g < l \leq d. \quad (16)$$

The entries of coefficient vectors of w and z can be sampled for full-enumeration as follows [5]:

$$w_i \leftarrow (-1)^{\lfloor \text{rand}_{[0..2]} \rfloor} \cdot \sqrt{\frac{\omega_i \cdot (\|v\|^2 - \|b_g^*\|^2)}{\|b_i^*\|^2 \cdot (\sum_{t=1}^{g-1} \omega_t)}}, \quad \text{where } \omega_i \leftarrow \text{Gamma}(1/2, 2). \quad (17)$$

$$z_{d-i+1} \leftarrow (-1)^{\lfloor \text{rand}_{[0..2]} \rfloor} \cdot \sqrt{\frac{\omega_i \cdot (\|v\|^2 - \|b_g^*\|^2)}{\sum_{t=1}^{g-1} \omega_t}}, \quad \text{where } \omega_i \leftarrow \text{Gamma}(1/2, 2). \quad (18)$$

3. Our Contribution

As mentioned, GNR enumeration in BKZ algorithm over a lattice block $\mathcal{L}_{[b_j, \dots, b_k]}$ returns the coefficient vector y , which can be used to compute the solution vector v by linear combination of this lattice block vectors as $v = \sum_{l=j}^k y_l b_l$.

To the best of our knowledge, our work in [5] is first precise analysis of sampling the coefficient vector y . As mentioned, our sampling methods for coefficient vectors is developed in three versions in [5], as follows:

- Our sampling method 1 (which is introduced in Claim 1 and Lemma 11 from [5]; see the pseudo-code in Algorithm 5 from [5]);
- Our sampling method 2 (see the pseudo-code in Algorithm 6 from [5]);
- Our sampling method 3 (see the pseudo-code in Algorithm 6 from [5]);

This paper proposes Algorithm 1 as the skeleton of our sampling method of coefficient vectors.

Note: For simplicity, Algorithm 1 assumes that cutting point is $\text{Cut} = d$, while in practical purposes, if $\text{Cut} < d$ then this is only needed to change the block of $\mathcal{L}_{[1,d]}$ with block size of d into the block of $\mathcal{L}_{[1,\text{Cut}]}$ with block size of Cut .

Algorithm 1 Sample_of w

Input: GSO norms of block given as $B_{[1,d]}^* = [\|b_1^*\| \dots \|b_d^*\|]$ at stage j , any type of bounding function $\mathcal{R}$, enumeration radii R , Sampled norm of enumeration solution as ℓ_1 , last nonzero index of $\mathcal{G}$, /* boolean parameter of RANDBLOCK */ GSO Coefficient Matrix of μ .

```

1:  $[\ell_1^2, \ell_2^2, \dots, \ell_{\mathcal{G}}^2] \leftarrow \text{Sample\_of\_}\ell_t^2(B_{[1,d]}^*, \mathcal{R}, R, \ell_1, \mathcal{G});$ 
2: for  $(l = \mathcal{G} + 1; l \leq d; l++)$   $y_l = 0;$ 
3:  $y_{\mathcal{G}} = 1;$ 
4: for  $(l = \mathcal{G} - 1; l \geq 1; l--)$  {
5:    $y_l = \frac{\sqrt{\ell_l^2 - \ell_{l+1}^2}}{\|b_l^*\|} - \sum_{i=l+1}^{\mathcal{G}} y_i \mu_{i,l} ; // \ell_x^2 = \|\pi_x(v)\|^2$ 
6: } //end for
7: for  $(l = 1; l \leq d; l++)$  {
8:    $w_l = y_l + \sum_{i=l+1}^d y_i \mu_{i,l} ;$ 
9:    $z_{d-l+1} = w_l \|b_l^*\|;$ 
10: } //end for

```

Output: Sampled coefficient vectors of w, z and y .

Lemma 1. For a GNR-enumeration over GSO block $[b_1^*, \dots, b_d^*]$ by any type of bounding function $\mathcal{R}$, enumeration radii R and current GSO Coefficient Matrix of μ , if there is a solution which is returned with sampled norm of ℓ_1 and last nonzero index of $\mathcal{G}$, Algorithm 1 gives the samples of coefficient vectors of w, z and y for this sampled solution based on Heuristic 2, just if the function “Sample_of_ ℓ_t^2 ” truly samples $\|\pi_t(v)\|$ for $1 \leq t \leq \mathcal{G}$ based on Heuristic 2.

See proof in Appendix C.

As mentioned in Lemma 1, Algorithm 1 gives the samples of coefficient vectors of w, z and y for this sampled solution based on Heuristic 2, just if the function of “Sample_of_ ℓ_t^2 ” truly samples the norm of $\pi_t(v)$ for $1 \leq t \leq \mathcal{G}$ based on Heuristic 2. Algorithm 3 in Appendix A uses Lemma 3 in [5] for implementing the function of “Sample_of_ ℓ_t^2 ”. Lemma 2 declares that this idea in Algorithm 3 (which is named as “Our Sampling method 4”) is not consistent with Heuristic 2 (i.e., it is impossible to use Lemma 3 in [5] for implementing the function of Sample_of_ ℓ_t^2).

Lemma 2. Under Heuristic 2, if the sampled solution of $v = \pi_1(v)$ would be projected on $\text{span}(b_1, \dots, b_t)^\perp$, then $\pi_t(v)$ cannot be considered as an independent uniformly distributed sampled vector on the projected lattice block of $[\pi_2(b_t), \dots, \pi_2(b_d)]$, and consequently the norm of $\pi_t(v)$ cannot be sampled by using Lemma 3 in [5] in the way which is introduced in Algorithm 3 (with corresponding parameters which are updated for this projected block in lines 3-10 from Algorithm 3 in Appendix A).

Proof. This proof is done by introducing a Counter-example. Lets assume a GSO basis of $B_{[1,2]}^* = [b_1^* = (1,0), b_2^* = (0,1)]$. For enumeration radius R and full-enumeration bounding function (i.e., vector of $\mathcal{R} = [1,1]$), if a solution vector v would be sampled uniformly from these lattice vectors in the circle of radius R based on Heuristic 2, then is the distribution of projected vector of $\pi_2(v)$ equal to uniform distribution over vertical vertex with basis of $[b_2^* = (0,1)]$? The answer is clearly “NO”. In fact, the distribution of projected vector of $\pi_2(v)$ has most frequency in point of $(0,0)$ from vertical vertex, while the least frequency can be observed in points of $(0, \pm[R])$ from vertical vertex, this means that the distribution of the projected vector of $\pi_2(v)$ in this example is not uniform. In other side, Lemma 3 in [5] is mainly defined based on uniformly distribution of lattice vectors in corresponding cylinder-intersection of the bounding function, so Lemma 3 in [5] cannot be used as “Sample_of_ ℓ_t^2 ” for the projected vector $\pi_i(v)$ where $i > 1$.

Our results Test 1 from Section 4 verifies Lemma 2 (see Fig.2).

For best accuracy in sampling coefficient vectors, this paper introduces Algorithm 2 as “Our sampling method 5”:

Algorithm 2 Sample_of_ℓ_t²_Ver5 (Our Sampling Method 5)

Input: GSO norms of block given as $B_{[1,d]}^* = [b_1^*, \dots, b_d^*]$, bounding function $\mathcal{R}$, enumeration radii R , Sampled norm of enumeration solution as ℓ_1 , last nonzero index of $\mathcal{g}$.

- 1: $\mathcal{R}_{UP1} \leftarrow \text{Get_UpperBound_of_}\mathcal{R}_1(d, \mathcal{R});$
- 2: $\mathcal{R}_{UP2} \leftarrow \text{Get_UpperBound_of_}\mathcal{R}_2(d, \mathcal{R}, \mathcal{R}_{UP1});$
- 3: $\mathcal{R}_{LOW} \leftarrow \text{Get_LowerBound_of_}\mathcal{R}(d, \mathcal{R}, \mathcal{g});$
- 4: $[E[\ell_1^2], \dots, E[\ell_d^2]] \leftarrow \text{Get_ExpectedValues_of_}\ell_t^2(B_{[1,d]}^*, \mathcal{R}, R, \ell_1, \mathcal{g}, \mathcal{R}_{LOW}, \mathcal{R}_{UP1}, \mathcal{R}_{UP2});$
- 5: $[\mathcal{V}[\ell_1^2], \dots, \mathcal{V}[\ell_d^2]] \leftarrow \text{Get_Variances_of_}\ell_t^2(B_{[1,d]}^*, \mathcal{R}, R, \ell_1, [E[\ell_1^2], \dots, E[\ell_d^2]]);$
- 6: $[\ell_1^2, \dots, \ell_d^2] \leftarrow \text{Get_Sample_of_}\ell_t^2(B_{[1,d]}^*, \ell_1, \mathcal{g}, [E[\ell_1^2], \dots, E[\ell_d^2]], [\mathcal{V}[\ell_1^2], \dots, \mathcal{V}[\ell_d^2]]);$

Output: Sampled norms of $\ell_1^2, \ell_2^2, \dots, \ell_d^2$.

Algorithm 2 includes 5 main steps which are defined as follows.

Step 1. Computing Upper-bound bounding functions:

This phase (lines 1-2 in Algorithm 2) defines a Step bounding function of $\mathcal{R}_{UP1}$ and a modification of $\mathcal{R}_{UP1}$ as $\mathcal{R}_{UP2}$ including the entries which are closely lower-bounded by input bounding function $\mathcal{R}$. See the pseudo-codes of Get_UpperBound_of_ℛ_1 and Get_UpperBound_of_ℛ_2 respectively in Algorithm 4 and Algorithm 5 from Appendix B. Fig.3 shows the upper-bound bounding functions of $\mathcal{R}_{UP1}$ and $\mathcal{R}_{UP2}$ for input bounding function of “BF2 [Psucc=0.0002]”.

Step 2. Computing Lower-bound bounding functions:

This phase (lines 3 in Algorithm 2) defines a piecewise-linear bounding function of $\mathcal{R}_{LOW}$ including the entries which are closely upper-bounded by input bounding function $\mathcal{R}$. See the pseudo-codes of Get_LowerBound_of_ℛ in Algorithm 6 from Appendix B. Fig.3 shows the lower-bound bounding function of $\mathcal{R}_{LOW}$ for an input bounding function of “BF2 [Psucc=0.0002]”.

Step 3. Computing the expected value of ℓ_t^2 :

In this step, the expected value of $\ell_{d-i+1}^2 = \|\pi_{d-i+1}(v)\|^2$ can be computed as follows.

$$X_i = \sum_{t=1}^i z_t^2 / R^2, \text{ for } 1 \leq i \leq d \Rightarrow$$

$$R^2 E[X_i] = R^2 E\left[\sum_{t=1}^i z_t^2 / R^2\right] = E\left[\sum_{t=1}^i z_t^2\right] = E[z_1^2 + \dots + z_{d-g+1}^2 + z_{d-g+2}^2 + \dots + z_i^2] = E[z_{d-g+1}^2 + z_{d-g+2}^2 + \dots + z_i^2] \Rightarrow$$

By using (13), Theorem 2 in [5] and (32):

$$R^2 E[X_i] = E[\|\pi_{d-i+1}(v)\|^2] = E[\ell_{d-i+1}^2]. \quad (19)$$

By using Claim 1, the expected value of $R^2 X_i = \sum_{t=1}^i z_t^2$ for Step bounding function $\mathcal{R}_{Step}$ can be approximated.

Claim 1. Under Heuristic 2, Algorithm 8 closely approximates the expected value of $\ell_{d-i+1}^2 = \|\pi_{d-i+1}(v)\|^2 = R^2 X_i = \sum_{t=1}^i z_t^2$ for Step bounding function $\mathcal{R}_{Step}$.

To have better sense, see the shape of $E[\ell_{d-i+1}^2]$ for Step bounding function of “BF6 [Psucc=0.000147]” by our results in Fig.4. Also, by using Claim 2, the expected value of $R^2 X_i = \sum_{t=1}^i z_t^2$ for Piecewise-linear bounding function $\mathcal{R}_{Piece}$ can be approximated.

Claim 2. Under Heuristic 2, Algorithm 9 closely approximates the expected value of $\ell_{d-i+1}^2 = \|\pi_{d-i+1}(v)\|^2 = R^2 X_i = \sum_{t=1}^i z_t^2$ for Piecewise-linear bounding function $\mathcal{R}_{Piece}$.

To have better sense, see the shape of $E[\ell_{d-i+1}^2]$ for Piecewise-linear bounding function of “BF5 [Psucc=0.00000011]” by our results in Fig.5. Finally, by using Claim 3, the expected value of $R^2 X_i = \sum_{t=1}^i z_t^2$ for any shape of bounding function $\mathcal{R}$ can be approximated.

Claim 3. Under Heuristic 2, for any shape of bounding function $\mathcal{R}$, if $\frac{\partial(\mathcal{R}[i])}{\partial i} < \frac{2}{d}$ then Algorithm 7 closely approximates the expected value of $\ell_{d-i+1}^2 = \|\pi_{d-i+1}(v)\|^2 = R^2 X_i = \sum_{t=1}^i z_t^2$.

To have better sense, see the shape of $E[\ell_t^2]$ for bounding function of “BF2 [Psucc=0.0002]” by our results in Fig.6.

Step 4. Computing the variance of ℓ_t^2 :

In this step, the variance of $\ell_{d-i+1}^2 = \|\pi_{d-i+1}(v)\|^2$ can be computed as follows:

$$R^4 \mathcal{V}[X_i] = \mathcal{V}[\|\pi_{d-i+1}(v)\|^2] = \mathcal{V}[\ell_{d-i+1}^2]. \quad (20)$$

To approximate the variance of ℓ_{d-i+1}^2 for an input bounding function $\mathcal{R}$, at first, this paper estimates the variance of ℓ_{d-i+1}^2 for full-enumeration bounding function, then update this estimation for any arbitrary bounding function $\mathcal{R}$. The variance of $X_i = \sum_{t=1}^i z_t^2 / R^2$ for full-enumeration can be approximated by following lemma:

Lemma 3. Under Heuristic 2, if random variable $w_l^2 = z_l^2 / \|b_l^*\|^2$ (for $l = d - i + 1$) would be sampled by (17) under full-enumeration with solution norm of $\|v\|$, the variance of $X_i = \sum_{t=1}^i z_t^2 / R^2$ can be approximated by (21):

$$\mathcal{V}[X_i]_{\text{full}} = \left(\frac{\|v\|^2 - \|b_d^*\|^2}{R^2} \right)^2 \frac{2id - 2i^2 - 2d + 2i}{d^3 - d^2 - d + 1}. \quad (21)$$

See proof in Appendix D.

Corollary 1. By Lemma 3 and (20), the variance of ℓ_{d-i+1}^2 for full-enumeration can be computed as follows:

$$R^4 \mathcal{V}[X_i]_{\text{full}} = \mathcal{V}[\|\pi_{d-i+1}(v)\|^2]_{\text{full}} = \mathcal{V}[\ell_{d-i+1}^2]_{\text{full}} = (\|v\|^2 - \|b_d^*\|^2)^2 \frac{2id - 2i^2 - 2d + 2i}{d^3 - d^2 - d + 1}. \quad (22)$$

The variance of $X_i = \sum_{t=1}^i z_t^2 / R^2$ for an arbitrary bounding function $\mathcal{R}$ can be estimated by Approximation 1:

Approximation 1. Under Heuristic 2, for an arbitrary bounding function $\mathcal{R}$, if the variance of $X_i = \sum_{t=1}^i z_t^2 / R^2$ for full-enumeration is approximated as $\mathcal{V}[X_i]_{\text{full}}$ and the expected value of X_i is approximated as $E[X_i]$, then the variance of X_i for $\mathcal{R}$ can be approximated by (23):

$$\mathcal{V}[X_i] \leftarrow \mathcal{V}[X_i]_{\text{full}} \times \min \left(\left(\frac{\mathcal{R}^2[i] - E[X_i]}{\mathcal{R}_{\text{UP3}}^2[i] - E[X_i]_{\text{full}}} \right)^2, 1 \right), \quad (23)$$

where $\mathcal{R}_{\text{UP3}}^2[i] = \min(0.7 \times (E[X_i]_{\text{full}} + 3 \times \sqrt{\mathcal{V}[X_i]_{\text{full}}}) + 0.3, 1)$ and $E[X_i]_{\text{full}}$ can be computed by using Algorithm 8 in Claim 1 (since full-enumeration bounding function is a step bounding function) or it can be computed by Lemma 9 in [5]. By Approximation 1 and (20), the variance of ℓ_{d-i+1}^2 for an input bounding function $\mathcal{R}$ can be estimated by (24):

$$\mathcal{V}[\ell_{d-i+1}^2] \leftarrow R^4 \mathcal{V}[X_i]. \quad (24)$$

The pseudo-code of computing the variance of ℓ_t^2 is introduced in Algorithm 10.

Step 5. Generating the samples of ℓ_t^2 :

Under Heuristic 2, Approximation 2 generates the samples of $\ell_{t=d-i+1}^2$ for an arbitrary bounding function $\mathcal{R}$:

Approximation 2. Under Heuristic 2, the values of $\ell_{t=d-i+1}^2 = \|\pi_{d-i+1}(v)\|^2 = R^2 X_i = \sum_{t=1}^i z_t^2$ for an arbitrary bounding function $\mathcal{R}$ can be sampled by following formula:

$$\ell_t^2 \leftarrow \begin{cases} 1: 0, & \text{for } t = d \text{ downto } g + 1, \\ 2: \|b_g^*\|^2, & \text{for } t = g, \\ 3: \min(\min(\mathcal{R}^2[d - t + 1] \times R^2, \|v\|^2), \ell_{t+1}^2 + \max(Y_t + \min(E[\ell_t^2] - \ell_{t+1}^2 - \varepsilon, 0), 0)), & \text{for } t = g - 1 \text{ downto } 2, \\ 4: \|v\|^2, & \text{for } t = 1, \end{cases} \quad (25)$$

where $\varepsilon = 1/d$, $\Delta E_t = \max(E[\ell_t^2] - \ell_{t+1}^2, \varepsilon)$ and $Y_t \leftarrow \text{Gamma} \left(k = \frac{(\Delta E_t)^2}{\mathcal{V}[\ell_t^2]}, \theta = \frac{\mathcal{V}[\ell_t^2]}{\Delta E_t} \right)$.

The pseudo-code of sampling by Approximation 2 is introduced in Algorithm 11. Also, our results in Test 6 from Section 4, verify the statistical exactness of ‘‘Our Sampling Method 5’’ including Approximation 2.

Table 1. The configuration of our tests in this paper

Specification	Value
Dimension (β or d)	200
Last non-zero index (g)	200
Determinant of $\mathcal{L}_{[1,\beta]}$	0.2058911321e45
$\text{GH}^2(\mathcal{L}_{[1,\beta]})$	33.5495
Radius Factor (r_{FAC}^2)	1.1
Radius R^2	36.9045
$\ b_1^*\ ^2$	45657810
$\ b_g^*\ ^2$	0.12004e-11
Count of solutions in Full-Enum $r_{\text{FAC}}^\beta / 2$	6890.31

4. Test Results

This section introduces sufficient test results to show the correctness, exactness and value of our claims and contributions in sampling the coefficient vector of GNR-enumeration solution. The configuration of our tests in this section uses the parameters in Table 1. Also the success probability of bounding functions in this paper is computed by using Algorithm 7 from paper [2]. Following notes are introduced for better study of our results in this section:

- Based on the definition of $\ell_{t=d-i+1}^2 = R^2 X_i$ and $X_i = \sum_{t=1}^i z_t^2 / R^2$, the expressions of “Sample_X_i by R” in Figure 3 in [5], Figure 4 in [5] and Figure 5 in [5] should be re-written as “Sample_ $\ell_{t=d-i+1}^2$ by R”.
- The entries of all bounding functions in the test results of this section are squared.
- The block size d in all results of this section is even.
- The expression of “Our Sampling Method 4” refers to Algorithm 1 by using Algorithm 3 as Sample_of_ ℓ_t^2 .
- The expression of “Our Sampling Method 5” refers to Algorithm 1 by using Algorithm 2 as Sample_of_ ℓ_t^2 .

Also, the shape of all bounding functions which are used in our tests are shown in Fig.1.

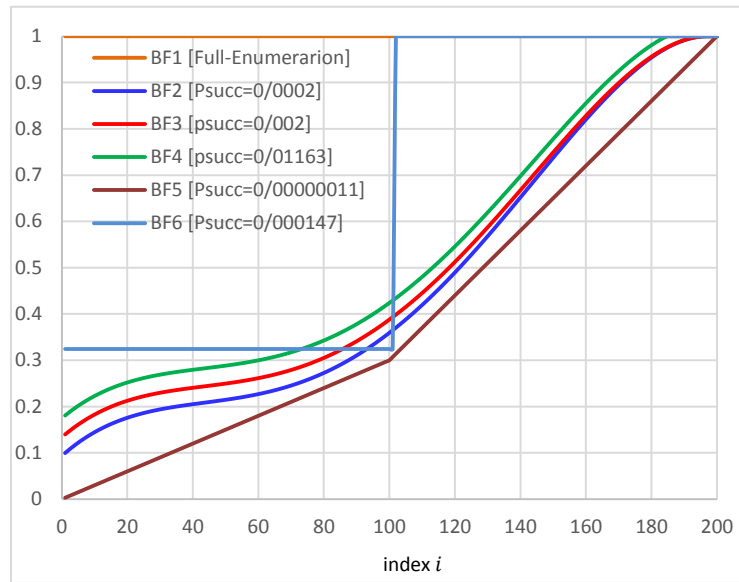


Fig.1. All bounding functions used in our tests results of this paper

- BF1 [Full-Enumeration]: It is a vector including the entries with values of 1 for full-enumeration.
- BF2 [Psucc=0.0002]: It is an arbitrary bounding function with success probability of $p_{succ} = 0.0002$.
- BF3 [Psucc=0.002]: It is an arbitrary bounding function with success probability of $p_{succ} = 0.002$.
- BF4 [Psucc=0.01163]: It is an arbitrary bounding function with success probability of $p_{succ} = 0.01163$.
- BF5 [Psucc=0.00000011]: It is a piecewise-linear bounding function with parameter of $a = 0.3$ and success probability of $p_{succ} = 0.00000011$.
- BF6 [Psucc=0.000147]: It is a Step bounding function with parameter of $a = 0.3244$ and success probability of $p_{succ} = 0.000147$.

Test 1. Rejecting the correctness of Algorithm 3 under Heuristic 2: In this test, by using the configuration in Table 1, our test results for 169 successful samples verify Lemma 2 which declares Algorithm 3 as “our sampling method 4” cannot be defined under Heuristic 2, since as shown in Fig.2, the expected value of samples by original sampling method are not closing to the expected value of samples by our sampling method 4. In fact, in this paper, this is tried to introduce some sampling method which generates some samples of coefficient vectors in the way that the statistical measurements of entries of these samples (e.g., expected value) are close to the corresponding measurements of entries of samples by original sampling method by (18), while Algorithm 3 as “our sampling method 4” dose not satisfy this goal at all!

Note: All figures in this paper which shows the shape of $\ell_{t=d-i+1}^2$ or $E[\ell_{t=d-i+1}^2]$ for some input bounding function $\mathcal{R}$, also shows the scaled entries of these bounding functions as $\mathcal{R}^2[i] \times R^2$ (such as: Fig.2, Fig.4, Fig.5, Fig.6, Fig.7 and Column 1 in Fig.8), while Fig.1 and Fig.3 just shows $\mathcal{R}^2[i]$.

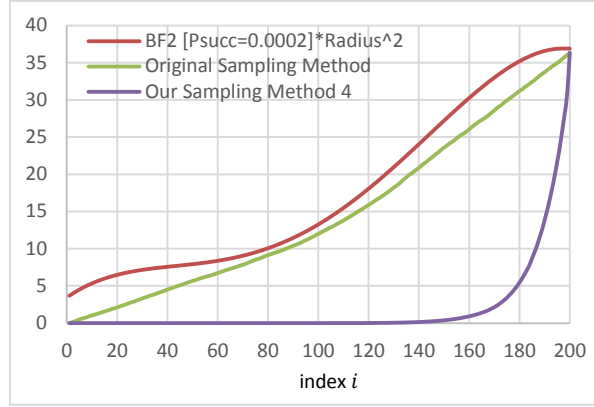


Fig.2. Comparison of expected value of entries of samples by original sampling method by (18) as $E[\ell_{t=d-i+1}^2]$ and our sampling method 4 by Algorithm 3 for bounding function of BF2 [Psucc=0.0002]

Test 2. Some Observations for Step 1, Step 2 and Step 3 in Our Sampling Method 5: In this test, by using the configuration in Table 1, Fig.3 shows the upper-bound bounding functions of $\mathcal{R}_{UP1}$ (as explained in Step 1 in Section 3) and $\mathcal{R}_{UP2}$ (as explained in Step 1 in Section 3) and lower-bound bounding function of $\mathcal{R}_{LOW}$ (as explained in Step 2 in Section 3) for an input bounding function “BF2 [Psucc=0.0002]”.

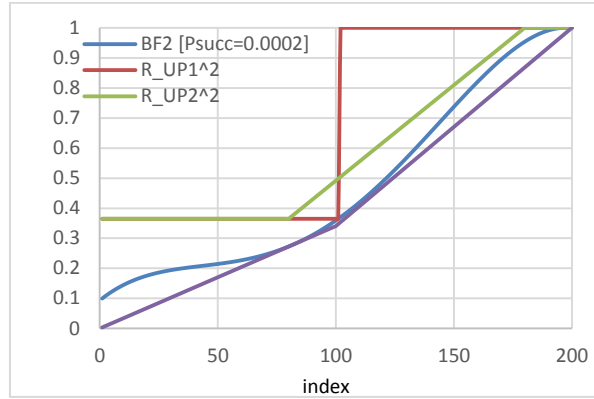
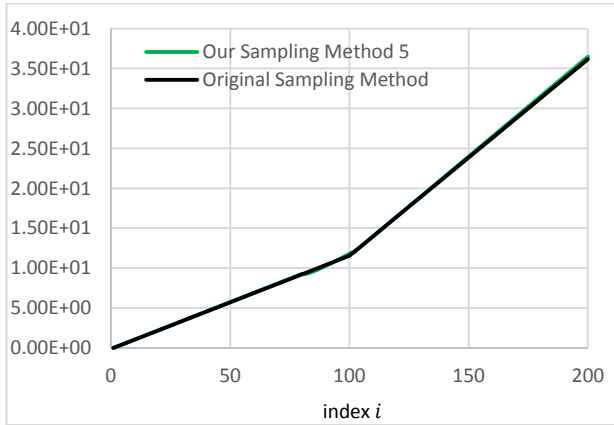
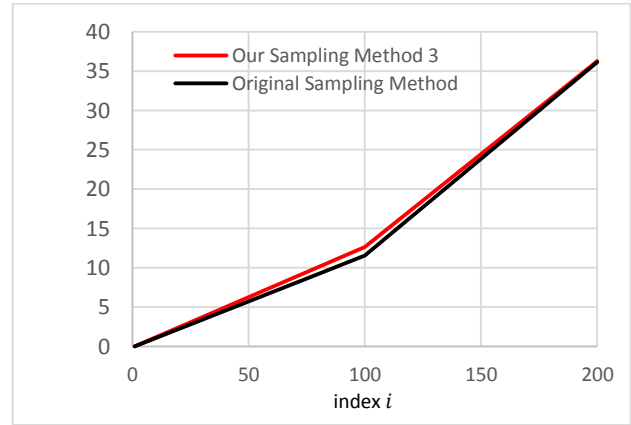


Fig.3. Upper-bound and Lower-bound bounding functions for input bounding function of “BF2 [Psucc=0.0002]”

Note: As mentioned, the entries of all bounding functions in the test results of this section (including “BF2 [Psucc=0.0002]” in Fig.3) are squared.



(1) Comparison of samples of “Original Sampling Method” and “Our Sampling Method 5”



(2) Comparison of samples of “Original Sampling Method” and “Our Sampling Method 3”

Fig.4. Comparison of $E[\ell_{t=d-i+1}^2]$ as the expected value of entries of samples for Original Sampling Method by (18) and Our Sampling Method 5 by Algorithm 2 and Our Sampling Method 3 for Step bounding function of “BF6 [Psucc=0.000147]”

Test 3. Verification of Claim 1: In this test, by using the configuration in Table 1, our test results for nearly 20000 successful samples verify the exactness of Claim 1, since as shown in Fig.4, the expected value of samples by original sampling method are more close to the expected value of samples by our sampling method 5 than the one by our sampling method 3. In fact, as shown in Fig.4, Claim 1 can be assumed as a strong assumption.

Test 4. Verification of Claim 2: In this test, by using the configuration in Table 1, our test results for nearly 20000 successful samples verify the exactness of Claim 2, since as shown in Fig.5, the expected value of samples by original sampling method are more close to the expected value of samples by our sampling method 5 than the one by our sampling method 3. However, as shown in Fig.5, Claim 2 is not as exact as Claim 1 (see Fig.4), this claim (Claim 2) can be assumed as a better assumption than using “Our sampling method 3”!

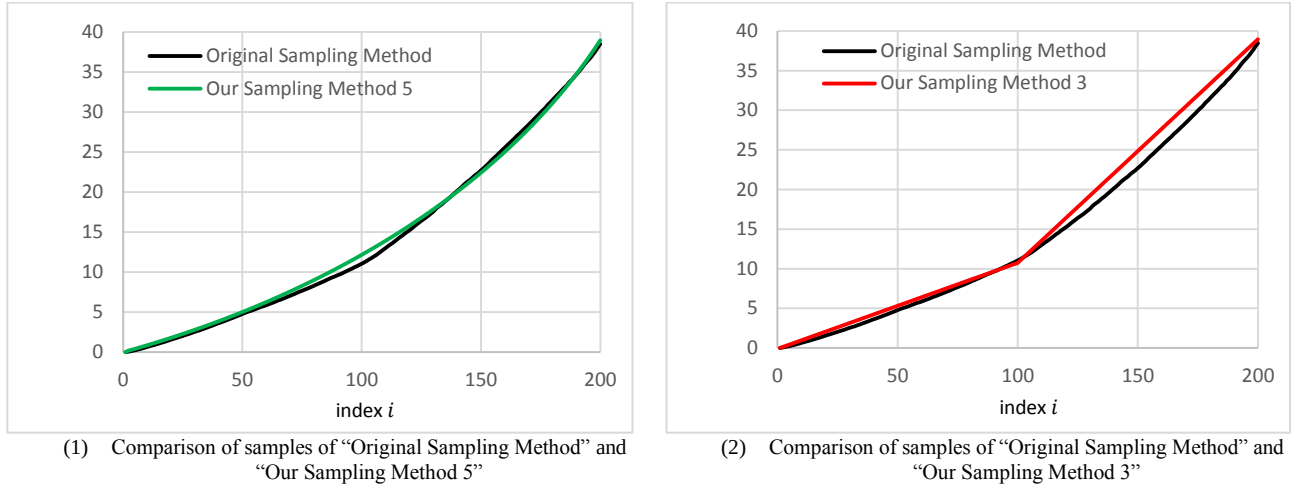


Fig.5. Comparison of $E[l_{t=d-i+1}^2]$ as the expected value of entries of samples for Original Sampling Method by (18) and Our Sampling Method 5 by Algorithm 3 and Our Sampling Method 3 for Piecewise-linear bounding function of “BF5 [Psucc=0.00000011]”

Test 5. Some Observations for Claim 3: Fig.6 shows that how the estimation of $E[l_{t=d-i+1}^2]$ in “Our Sampling Method 5” can be drawn between the estimations of $E[l_{t=d-i+1}^2]$ for $\mathcal{R}_{UP1}$ and $\mathcal{R}_{LOW}$. The sufficient comparison of the estimation of $E[l_{t=d-i+1}^2]$ in “Our Sampling Method 5” with “Our Sampling Method 3” is introduced in Test 6.

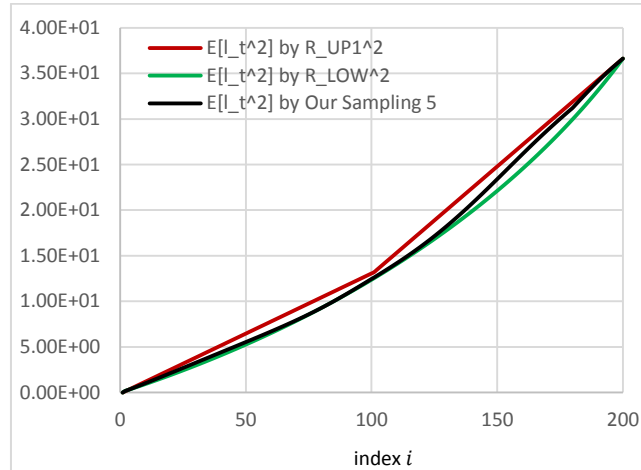


Fig.6. Estimation of $E[l_{t=d-i+1}^2]$ as the expected value of entries of samples for in “Our Sampling Method 5” between the estimations of $E[l_{t=d-i+1}^2]$ for $\mathcal{R}_{UP1}$ and $\mathcal{R}_{LOW}$ for input bounding function of “BF2 [Psucc=0.0002]”

Test 6. Final Comparison of “Our Sampling Method 5” and “Our Sampling Method 3”: In this test, by using the configuration in Table 1, we introduce our main (and final) comparison of “Our Sampling Method 5” and “Our Sampling Method 3” based on statistical measurements. Fig.7 shows the shapes of 20 samples of $\ell_{t=d-i+1}^2$ for input bounding function of “BF2 [Psucc=0.0002]” which are generated by “Original Sampling Method”, “Our Sampling Method 5” and “Our Sampling Method 3”. Based on Fig.7, this is shown that the samples of “Original Sampling Method” are more close to the samples of “Our Sampling Method 5” than the one by “Our Sampling Method 3”. Also Fig.8 introduces our best comparison between the samples generated by “Original Sampling Method”, “Our Sampling Method 5” and “Our Sampling Method 3”. This test uses nearly 100 successful samples for each methods in Fig.8.

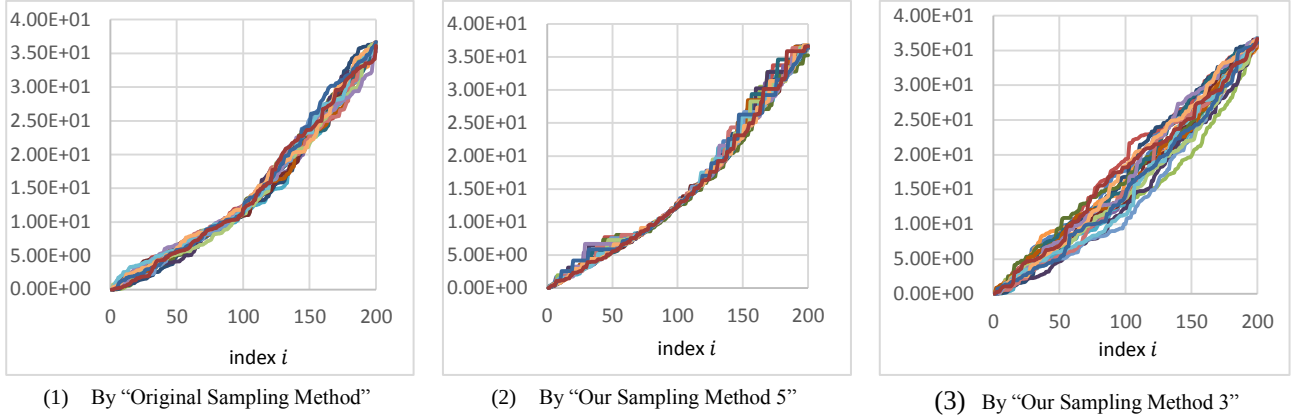
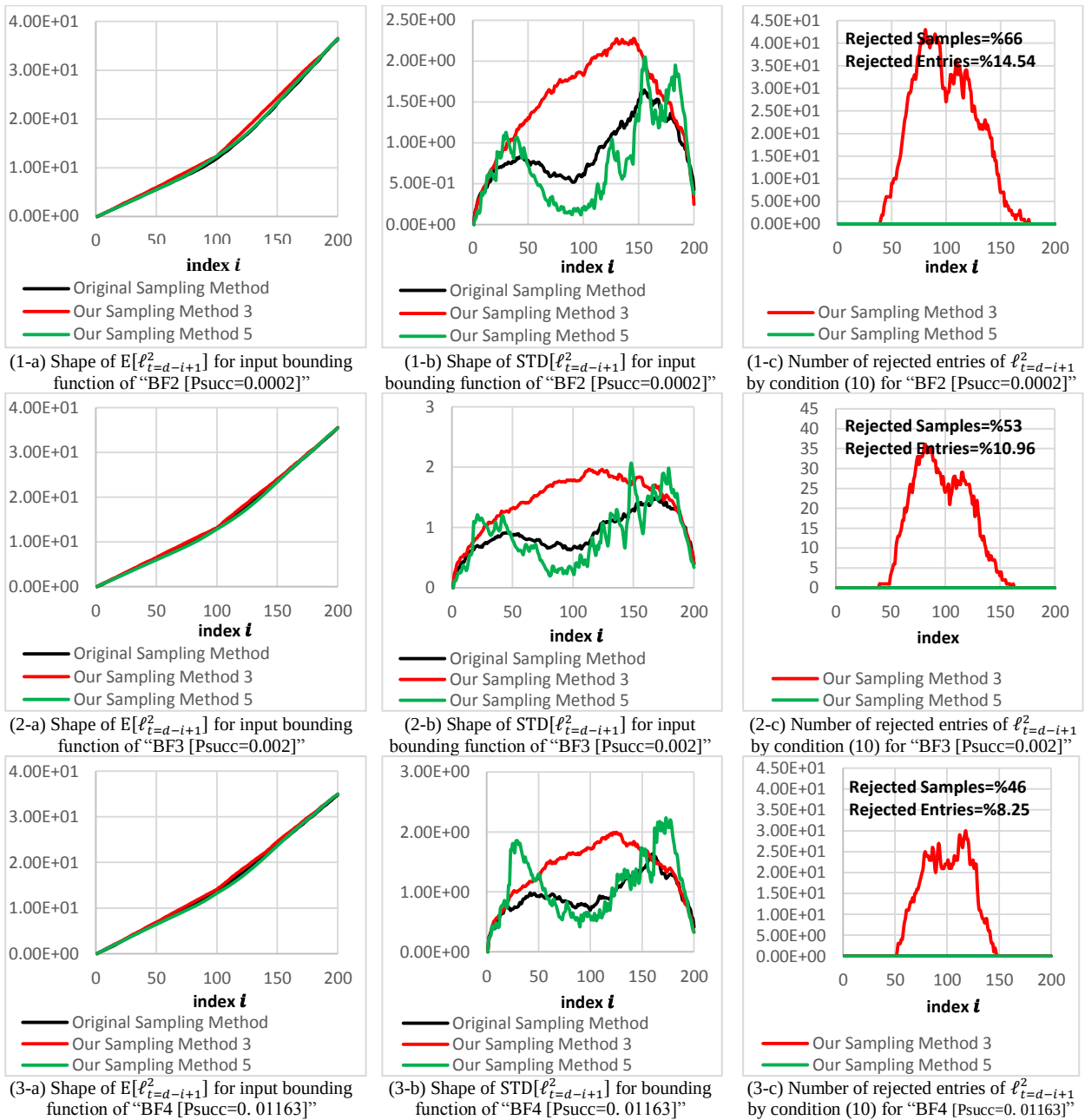

 Fig.7. The shape of 20 samples of $\ell^2_{t=d-i+1}$ for input bounding function of "BF2 [Psucc=0.0002]"

 Fig.8. Statistical specifications of nearly 100 assumed successful samples of coefficient vector ℓ or z for three input bounding functions by methods of "Original Sampling Method", "Our Sampling Method 3" and "Our Sampling Method 5"

Fig.8 shows the superiorities of “Our Sampling Method 5” over “Our Sampling Method 3” based on statistical measurements:

- The column 1 in Fig.8 shows that the expected value of the entries of the samples by “Original Sampling Method” are more close to the ones by “Our Sampling Method 5” than the ones by “Our Sampling Method 3”, however Fig.4 and Fig.5 introduce better comparisons for showing this superiority of “Our Sampling Method 5” over “Our Sampling Method 3”.
- The column 2 in Fig.8 shows that the standard deviation of the entries of the samples by “Original Sampling Method” are more close to the ones by “Our Sampling Method 5” than the ones by “Our Sampling Method 3”.
- The column 3 in Fig.8 shows the surprising number of violations of main condition of (10) by each entry of the coefficient vector samples of $\ell_{t=d-i+1}^2$ or z_i^2 in “Our Sampling Method 3”, while the number of violations of main condition of (10) by each entry of the samples by “Our Sampling Method 5” are strictly zero! Also, the total number of rejected samples and total number of violations of main condition of (10) by each entry of the samples in “Our Sampling Method 3” are written on the corresponding figures in column 3 of Fig.8.

5. Conclusion

For predicting the manner of BKZ in higher block sizes, practical running is not the way, therefore BKZ-simulations are introduced. One of the main components in design of BKZ-simulation with GNR enumeration is sampling the enumeration solution v . Our idea in Lemma 3 from [5] introduces a suitable method for sampling the norm of enumeration solution v , also first precise and explicit analysis of sampling the coefficient vector y is introduced in [5]. Our idea in Lemma 3 from [5] is accepted in this paper, but our sampling methods of 1, 2 and 3 cannot be accepted as exact sampling methods of coefficient vectors and would be revised in this paper.

Based on the test results of this paper and [5], if the success probability of bounding function would be much small or enumeration radius tends to the solution norm (i.e., $R \approx \|v\|$), the proposed sampling method 1, 2 and 3 would be less accurate. A new idea for approximating the expected value of coefficient vector entries of enumeration solution is introduced which is more close to the expected value of corresponding entries in basic rejection sampling in Algorithm 6 from [2]. There is no analysis on variance of each entry in coefficient vector of enumeration solution in [5], while our test results in this paper show disappointing variance for these three sampling methods (our sampling method 1, 2, 3)! Also a novel idea for approximating this variance is introduced which is more close to the variance of corresponding entries in basic rejection sampling in Algorithm 6 from [2]. Moreover there is no analysis on the number of violations from main condition of bounding function at each entry in coefficient vector of enumeration solution in [5] (see this main condition in relation (10)), while our test results in this paper show surprising much number of violations for these three sampling methods (our sampling method 1, 2, 3)! For example, for some arbitrary bounding functions with success probability of $p_{succ} = 0.011$ up to $p_{succ} = 0.0002$, the number of rejected samples would be 46% up to 66%. By using these new achievements and analysis, our sampling method 1, 2 and 3 (for sampling coefficient vectors) are revised. Our new revised method is not a rejection sampling, this means that, it includes no violation from main condition of GNR bounding function at each entry of coefficient vectors (see this main condition in relation (10)). Also the expected value and variance of each entry of coefficient vectors by our new sampling method are nearly more close to our new exact approximation of expected value and variance of corresponding entries in Algorithm 6 from [2] than the entries of coefficient vectors by our sampling method 1, 2, 3.

In fact, for designing a BKZ-simulation, even our novel idea for approximating the expected value of corresponding entries in Algorithm 6 from [2] is sufficiently exact (instead of sampling by “Our Sampling Method 5” in this paper), while its variance and number of violations at each entry are absolutely zero. In other words, using our new sampling method just tries to inject more reasonable randomness to the BKZ-simulation. However this work introduces some acceptable sampling method for coefficient vectors of GNR-enumeration, more efficient and more exact samplers are expected to be studied and introduced in future researches.

Appendix A. “Our Sampling Method 4”

Algorithm 3 shows the pseudo-code of “Our Sampling Method 4”.

Algorithm 3 Sample_of_ ℓ^2 _Ver4 (Our Sampling Method 4)

Input: GSO norms of block given as $B_{[1,d]}^* = [b_1^*, \dots, b_d^*]$ at stage j , any type of bounding function $\mathcal{R}$, enumeration radii R ,

Sampled norm of enumeration solution as ℓ_{new}^2 , last nonzero index of $\mathcal{G}$.

```

1:   $\ell_1^2 = \ell_{new}^2$ ;
2:  for( $i = 2; i \leq \mathcal{G} - 1; i++$ ){
3:       $gh = GH(\pi_i(\mathcal{L}_{[i,d]}))$ ;
4:       $R_{new} = \min(R \times \mathcal{R}_{d-i+1}, \ell_{i-1}^2)$ ;
5:       $r_{FAC} = \frac{R_{new}}{gh}$ ;
6:      for( $l = i; l \leq d; l++$ ){ $\mathcal{R}_{d-l+1} \leftarrow \frac{\min(R \times \mathcal{R}_{d-l+1}, \ell_{l-1}^2)}{R_{new}}$ };
7:       $\mathcal{R}_{new} = [\mathcal{R}_1, \dots, \mathcal{R}_{d-i+1}]$ ; //  $\mathcal{R}_{new}$  is a new bounding function with length of  $d - i + 1$ 
8:       $R \leftarrow R_{new}$ ;
9:       $p = p_{succ}(\mathcal{R}_{new})$ ;
10:      $V = \text{Determinant}([\pi_i(b_i), \pi_i(b_{i+1}), \dots, \pi_i(b_d)])$ ;
11:      $\ell_i^2 = \text{Sample}_{\|v\|}(d - i + 1, p, gh, r_{FAC}, V, v_{d-i+1})$ ; //  $v_{d-i+1}$  as volume of  $(d - i + 1)$ -dimensional ball with unit radius
12: } //end for
13:  $\ell_{\mathcal{G}}^2 = \|b_{\mathcal{G}}^*\|^2$ ;

```

Output: Sampled norms of $\ell_1^2, \ell_2^2, \dots, \ell_{\mathcal{G}}^2$.

Appendix B. Sub-functions of “Our Sampling Method 5”

The pseudo-codes of all sub-functions of “Our Sampling Method 5” are introduced in this appendix. As mentioned, the block size d in all algorithms of this appendix is assumed even. Algorithm 4 generates a Step bounding function including the entries which are closely lower-bounded by input bounding function $\mathcal{R}$.

Algorithm 4 Get_UpperBound_of_ $\mathcal{R}_1$

Input: Block size d , bounding function $\mathcal{R}$.

```

1:   $a = \mathcal{R}^2 \left\lfloor \frac{d}{2} \right\rfloor + 0.000001$ ;
2:  for( $i = 1; i \leq \frac{d}{2}; i++$ )  $\mathcal{R}_{UP1}^2[i] = a$ ;
3:  for( $i = \frac{d}{2} + 1; i \leq d; i++$ )  $\mathcal{R}_{UP1}^2[i] = 1$ ;

```

Output: Upper_bound bounding function $\mathcal{R}_{UP1}$.

Algorithm 5 generates a modification of upper-bound Step bounding function of $\mathcal{R}_{UP1}$, as $\mathcal{R}_{UP2}$, including the entries which are closely lower-bounded by input bounding function $\mathcal{R}$.

Algorithm 5 Get_UpperBound_of_ $\mathcal{R}_2$

Input: Block size d , bounding function $\mathcal{R}$, Upper_bound bounding function $\mathcal{R}_{UP1}$.

```

1:   $a = \mathcal{R}_{UP1}^2[1]$ ;
2:   $A = \frac{1-a}{0.5d}$ ;
3:  for( $i = 0.9d; i \geq 0.4d; i--$ ){ //begin for1
4:       $y = A(i - 0.4d) + a$ ;
5:      if( $y < \mathcal{R}_i^2$ )  $A \leftarrow \frac{\mathcal{R}_i^2 - a}{i - 0.4d}$ ; } //end for1
6:   $i = 1$ ;
7:  for( $i < 0.4d; i++$ )  $\mathcal{R}_{UP2}^2[i] = a$ ;
8:  for(; ) { //begin for2
9:       $y = A(i - 0.4d) + a$ ;
10:     if( $y \leq 1$ ) {  $\mathcal{R}_{UP2}^2[i] = y$ ;  $i++$ ; }
11:     else break; } //end for2
12:  for( $i \leq d; i++$ )  $\mathcal{R}_{UP2}^2[i] = 1$ ;

```

Output: Upper_bound bounding function $\mathcal{R}_{UP2}$.

Algorithm 6 defines a piecewise-linear bounding function including the entries which are closely upper-bounded by input bounding function $\mathcal{R}$.

Algorithm 6 Get_LowerBound_of_ℛ

Input: Block size d , bounding function $\mathcal{R}$, Last non_zero index of $\mathcal{g}$.

```

1:  a = 1;
2:  i = d - g + 1;
3:  for(; i ≤ d/2; i++){
4:      if( $\frac{2ia}{d} > \mathcal{R}^2[i]$ ) a =  $\frac{\mathcal{R}^2[i] \times d}{2i}$ ;
5:      for(i = d/2 + 1; i ≤ d; i++){
6:          if( $2a - 1 + \frac{2i(1-a)}{d} > \mathcal{R}^2[i]$ ) {
7:              a =  $\frac{\mathcal{R}_i^2 d + d - 2i}{2d - 2i}$ ;
8:          }
9:      }
10: for(; i ≤ d/2; i++){ $\mathcal{R}_{Low}^2[i] = 2ia/d$ ; }
10: for(; i ≤ d/2; i++){ $\mathcal{R}_{Low}^2[i] = 2a - 1 + 2i(1 - a)/d$ ; }

```

Output: Lower_bound bounding function $\mathcal{R}_{Low}$.

Algorithm 7 returns the expected value of ℓ_t^2 for an input bounding function $\mathcal{R}$ and lattice block $\mathcal{L}_{[1,d]}$.

Note: The notation of $E[\ell_t^2]$ shows the expected value of ℓ_t^2 , also $e\ell_i^2$ shows the value of $E[\ell_{d-i+1}^2]/\|v\|^2$ (i.e., the index of i in $e\ell_i^2$ and bounding functions $\mathcal{R}[i]$ is consistent, while $i = d - t + 1$ in $E[\ell_t^2]$).

Algorithm 7 Get_ExpectedValues_of_ℓ²

Input: GSO norms of block given as $B_{[1,d]}^* = [b_1^*, \dots, b_d^*]$, bounding function $\mathcal{R}$, enumeration radii R ,

Sampled norm of enumeration solution as ℓ_1 , last nonzero index of $\mathcal{g}$, bounding functions of $\mathcal{R}_{Low}$, $\mathcal{R}_{UP1}$ and $\mathcal{R}_{UP2}$.

```

1:   $[E[\ell_{up}^2], \dots, E[\ell_d^2]] \leftarrow \text{Get\_E}[\ell_t^2]_{\text{of\_}\mathcal{R}_{\text{Step}}}(B_{[1,d]}^*, \mathcal{R}_{UP1}, R, \ell_1, \mathcal{g})$ ;
2:   $[E[\ell_{low}^2], \dots, E[\ell_d^2]] \leftarrow \text{Get\_E}[\ell_t^2]_{\text{of\_}\mathcal{R}_{\text{Piece}}}(B_{[1,d]}^*, \mathcal{R}_{Low}, R, \ell_1, \mathcal{g})$ ;
3:  for(i = 1; i ≤ d; i++){
4:       $E[\ell_{d-i+1}^2] \leftarrow \frac{(\mathcal{R}^2[i] - \mathcal{R}_{Low}^2[i]) \times E[\ell_{d-i+1}^2]_{up} + (\mathcal{R}_{UP2}^2[i] - \mathcal{R}^2[i]) \times E[\ell_{d-i+1}^2]_{low}}{\mathcal{R}_{UP2}^2[i] - \mathcal{R}_{Low}^2[i]}$ ;

```

Output: $E[\ell_1^2], \dots, E[\ell_d^2]$ for bounding function $\mathcal{R}$.

Algorithm 8 introduces the pseudo-code of one of the main sub-function of Algorithm 7.

Algorithm 8 Get_ℓ²_{of_ℛ_{Step}}

Input: GSO norms of block given as $B_{[1,d]}^* = [b_1^*, \dots, b_d^*]$, Step bounding function $\mathcal{R}$, enumeration radii R ,

Sampled norm of enumeration solution as ℓ_1 , last nonzero index of $\mathcal{g}$.

```

1:  gh = GH( $\pi_{d/2+1}(\mathcal{L}_{[1,d]})$ );
2:   $r_{FAC} = R \times \mathcal{R}[d/2]/gh$ ;
3:   $E[\ell_{d/2+1}^2] = \min(\frac{R^2}{2}, R^2 \mathcal{R}^2[d/2](1 - \frac{4}{d}))$ ;
4:   $A = \frac{E[\ell_{d/2+1}^2] - \|b_g^*\|^2}{g - d/2 - 1}$ ;
5:  for(i = 1; i ≤ d/2; i++){
6:       $E[\ell_{d-i+1}^2] \leftarrow A \times (i - d + g - 1) + \|b_g^*\|^2$ ;
7:  }
8:   $A = \frac{\ell_1^2 - E[\ell_{d/2+1}^2]}{d/2}$ ;
9:  for(i = d/2 + 1; i ≤ d; i++){
10:      $E[\ell_{d-i+1}^2] \leftarrow A \times (i - d/2) + E[\ell_{d/2+1}^2]$ ;

```

Output: $E[\ell_1^2], \dots, E[\ell_d^2]$ for Step bounding function $\mathcal{R}$.

Algorithm 9 introduces the pseudo-code of other main sub-function of Algorithm 7.

Algorithm 9 Get_ℓ²_{of_ℛ_{Piece}}

Input: GSO norms of block given as $B_{[1,d]}^* = [b_1^*, \dots, b_d^*]$, Piecewise_linear bounding function $\mathcal{R}$, enumeration radii R ,

Sampled norm of enumeration solution as ℓ_1 , last nonzero index of $\mathcal{g}$.

```

1:   $p_0 = \frac{p_{succ}(\mathcal{R})}{d}$ ;
2:  step = 1/32;  $\epsilon = p_0/100$ ;
3:  for(p = 0; p < p_0 -  $\epsilon$ ; step = step × 2){//begin for1
4:      B = B + step;
5:       $A = B + \sqrt{2B^2 + 2B}$ ;
6:      for(i = 1; i ≤ d; i++){
7:           $e\ell_i^2 = 1 + A - \sqrt{(1 + A + B)^2 - (i/d + A)^2}$ ;
8:      }
9:       $e\ell_i^2 \leftarrow (e\ell_i^2 - e\ell_1^2) \times \frac{1 - \|b_g^*\|^2/\|v\|^2}{1 - e\ell_1^2} + \|b_g^*\|^2/\|v\|^2$ ;
10:     p =  $p_{succ}(e\ell^2)$ ; }//end for1

```

```

11:  step =  $\frac{\text{step}}{2}$ ; B = B - step; A = B +  $\sqrt{2B^2 + 2B}$ ;
12:  while(true){//begin while1
13:      if( $p_0 - \varepsilon \leq p \leq p_0 + \varepsilon$ ) break;
14:      for( $i = 1; i \leq d; i++$ ){
15:           $e\ell_i^2 = 1 + A - \sqrt{(1 + A + B)^2 - (i/d + A)^2}$ ;
16:          for( $i = 1; i \leq d; i++$ ){
17:               $e\ell_i^2 \leftarrow \left( (e\ell_i^2 - e\ell_1^2) \times \frac{1 - \|b_\phi^*\|^2 / \|v\|^2}{1 - e\ell_1^2} + \|b_\phi^*\|^2 / \|v\|^2 \right)$ ;
18:               $p = p_{\text{succ}}(e\ell^2)$ ;
19:              if( $p < p_0 - \varepsilon$ ){
20:                  step =  $\frac{\text{step}}{2}$ ; B = B + step; A = B +  $\sqrt{2B^2 + 2B}$ ;
21:                  if( $p_0 + \varepsilon < p$ ){
22:                      step =  $\frac{\text{step}}{2}$ ; B = B - step; A = B +  $\sqrt{2B^2 + 2B}$ ;
23:                  }
24:              }
25:          }
26:      }
27:  }
28:  for( $i = 1; i \leq d; i++$ )  $\{E[\ell_{d-i+1}^2] \leftarrow e\ell_i^2 \times \|v\|^2\}$ 

```

Output: $E[\ell_1^2], \dots, E[\ell_d^2]$ for Piecewise_linear bounding function $\mathcal{R}$.

Algorithm 10 returns the variance of ℓ_t^2 for an arbitrary bounding function $\mathcal{R}$ and lattice block $\mathcal{L}_{[1,d]}$.

Algorithm 10 Get_Variances_of_ ℓ_t^2

Input: GSO norms of block given as $B_{[1,d]}^* = [b_1^*, \dots, b_d^*]$, Step bounding function $\mathcal{R}$, enumeration radii R , Sampled norm of enumeration solution as $\ell_1, [E[\ell_1^2], \dots, E[\ell_d^2]]$ for bounding function $\mathcal{R}$.

```

1:   $\mathcal{R}_{\text{full}} \leftarrow [1, 1, \dots, 1]$ ;
2:   $[E[\ell_1^2]_{\text{full}}, \dots, E[\ell_d^2]_{\text{full}}] \leftarrow \text{Get\_E}[\ell_t^2]_{\text{of\_}\mathcal{R}_{\text{Step}}}(B_{[1,d]}^*, \mathcal{R}_{\text{full}}, R, \ell_1, \phi_0 = d)$ ;
3:  for( $i = 1; i \leq d; i++$ ){
4:       $\mathcal{V}[\ell_{d-i+1}^2]_{\text{full}} = (\|v\|^2 - \|b_d^*\|^2)^2 \frac{2id - 2i^2 - 2d + 2i}{d^3 - d^2 - d + 1}$ ;
5:       $\mathcal{R}_{\text{UP3}}^2[i] = \min\left(\frac{0.7}{R^2} (E[\ell_{d-i+1}^2]_{\text{full}} + 3 \times \sqrt{\mathcal{V}[\ell_{d-i+1}^2]_{\text{full}}}) + 0.3, 1\right)$ ;
6:       $\mathcal{V}[\ell_{d-i+1}^2] \leftarrow \mathcal{V}[\ell_{d-i+1}^2]_{\text{full}} \times \min\left(\left(\frac{R^2 \mathcal{R}^2[i] - E[\ell_{d-i+1}^2]}{R^2 \mathcal{R}_{\text{UP3}}^2[i] - E[\ell_{d-i+1}^2]_{\text{full}}}\right)^2, 1\right)$ 

```

Output: $[\mathcal{V}[\ell_1^2], \dots, \mathcal{V}[\ell_d^2]]$ for bounding function $\mathcal{R}$.

Algorithm 11 generates the samples of ℓ_t^2 in “Our Sampling Method 5” efficiently, for an arbitrary bounding function $\mathcal{R}$ and lattice block $\mathcal{L}_{[1,d]}$.

Algorithm 11 Get_Sample_of_ ℓ_t^2

Input: GSO norms of block given as $B_{[1,d]}^* = [b_1^*, \dots, b_d^*]$, Sampled norm of enumeration solution as ℓ_1 , last nonzero index of ϕ , enumeration radii R , $[E[\ell_1^2], \dots, E[\ell_d^2]]$ for bounding function $\mathcal{R}$, $[\mathcal{V}[\ell_1^2], \dots, \mathcal{V}[\ell_d^2]]$ for bounding function $\mathcal{R}$.

```

1:  for( $t = d; t \geq \phi + 1; t--$ )  $\ell_t^2 \leftarrow 0$ ;
2:   $\ell_\phi^2 \leftarrow \|b_\phi^*\|^2$ ;
3:   $\varepsilon = \ell_1^2 / d$ ;
4:  for( $t = \phi - 1; t \geq 2; t--$ ){
5:       $\Delta E_t = \max(E[\ell_t^2] - \ell_{t+1}^2, \varepsilon)$ ;
6:       $Y_t \leftarrow \text{Gamma}\left(k = \frac{(\Delta E_t)^2}{\mathcal{V}[\ell_t^2]}, \theta = \frac{\mathcal{V}[\ell_t^2]}{\Delta E_t}\right)$ ;
7:       $\ell_t^2 \leftarrow \min(\min(\mathcal{R}^2[d - t + 1] \times R^2, \|v\|^2), \ell_{t+1}^2 + \max(Y_t + \min(E[\ell_t^2] - \ell_{t+1}^2 - \varepsilon, 0), 0))$ ;
8:  }
9:   $\ell_1^2 \leftarrow \|v\|^2$ ;

```

Output: samples of $\ell_1^2, \dots, \ell_d^2$ for bounding function $\mathcal{R}$.

Appendix C. Proof of Lemma 1

If sampled solution of $v = \pi_1(v)$ would be projected on $\text{span}(b_1)^\perp$ as $\pi_2(v)$, the under-laying lattice block of $[\pi_1(b_1), \pi_1(b_2), \dots, \pi_1(b_d)]$ is modified into $[\pi_2(b_2), \pi_2(b_3), \dots, \pi_2(b_d)]$, then this projected solution of $v = \pi_1(v)$ can be re-written with coefficient vector y by using (12) as follows:

$$\pi_1(v) = (y_1, y_2, \dots, y_d) \begin{bmatrix} \pi_1(b_1) \\ \vdots \\ \pi_1(b_d) \end{bmatrix} = (y_1, y_2, \dots, y_d) \begin{bmatrix} b_1^* \\ \vdots \\ b_d^* + \sum_{t=1}^{d-1} \mu_{d,t} b_t^* \end{bmatrix}, \quad (26)$$

$$\pi_2(v) = (y_2, y_3, \dots, y_d) \begin{bmatrix} \pi_2(b_2) \\ \vdots \\ \pi_2(b_d) \end{bmatrix} = (y_2, y_3, \dots, y_d) \begin{bmatrix} b_2^* \\ \vdots \\ b_d^* + \sum_{t=2}^{d-1} \mu_{d,t} b_t^* \end{bmatrix}. \quad (27)$$

For more projection over $\text{span}(b_1, b_2, \dots)^\perp$, this projected solution can be re-written as follows:

$$\pi_l(v) = (y_l, y_{l+1}, \dots, y_d) \begin{bmatrix} \pi_l(b_l) \\ \vdots \\ \pi_l(b_d) \end{bmatrix} = (y_l, y_{l+1}, \dots, y_d) \begin{bmatrix} b_l^* \\ \vdots \\ b_d^* + \sum_{t=l}^{d-1} \mu_{d,t} b_t^* \end{bmatrix}, \quad (28)$$

$$\pi_g(v) = (y_g, y_{g+1}, \dots, y_d) \begin{bmatrix} \pi_g(b_g) \\ \vdots \\ \pi_g(b_d) \end{bmatrix} = (y_g, y_{g+1}, \dots, y_d) \begin{bmatrix} b_g^* \\ \vdots \\ b_d^* + \sum_{t=g}^{d-1} \mu_{d,t} b_t^* \end{bmatrix}. \quad (29)$$

where $y_g = 1, y_{g+1} = 0, \dots, y_d = 0$.

The norm of enumeration solution v can be expanded on the original lattice block of $[\pi_1(b_1), \dots, \pi_1(b_d)]$ as follows:

$$\begin{aligned} \|\pi_1(v)\|^2 &= (y_1 + \sum_{t=2}^d y_t \mu_{t,1})^2 \|b_1^*\|^2 + \dots + (y_g + \sum_{t=g+1}^d y_t \mu_{t,g})^2 \|b_g^*\|^2 + \dots + y_d^2 \|b_d^*\|^2 = \\ &= (y_1 + \sum_{t=2}^g y_t \mu_{t,1})^2 \|b_1^*\|^2 + \dots + (y_{g-1} + \mu_{g,g-1})^2 \|b_{g-1}^*\|^2 + \|b_g^*\|^2 = (y_1 + \sum_{t=2}^g y_t \mu_{t,1})^2 \|b_1^*\|^2 + \|\pi_2(v)\|^2, \end{aligned} \quad (30)$$

where $y_g = 1$ and $y_{g+1} = 0 \dots y_d = 0$.

Note: The sampled norm of $\|\pi_l(v)\|^2$ is shown by ℓ_l^2 in this paper (also $\|\pi_1(v)\|^2$ is shown by $\ell_{new}^2 = \ell_1^2$). The sampled norm of $\|\pi_2(v)\|^2$ shown by ℓ_2^2 in Algorithm 1, can be expanded as follows:

$$\begin{aligned} \|\pi_2(v)\|^2 &= (y_2 + \sum_{t=3}^g y_t \mu_{t,2})^2 \|b_2^*\|^2 + \dots + (y_{g-1} + \mu_{g,g-1})^2 \|b_{g-1}^*\|^2 + \|b_g^*\|^2 = \\ &= (y_2 + \sum_{t=3}^g y_t \mu_{t,2})^2 \|b_2^*\|^2 + \|\pi_3(v)\|^2. \end{aligned} \quad (31)$$

For more projection of $\pi_l(v)$ up to $\pi_g(v)$, the norm of projected solution can be expanded as follows (note: the sampled norm of $\|\pi_l(v)\|^2$ is shown by ℓ_l^2 in this paper):

$$\begin{aligned} \|\pi_l(v)\|^2 &= (y_l + \sum_{t=l+1}^g y_t \mu_{t,l})^2 \|b_l^*\|^2 + \dots + (y_{g-1} + \mu_{g,g-1})^2 \|b_{g-1}^*\|^2 + \|b_g^*\|^2 = \\ &= (y_l + \sum_{t=l+1}^g y_t \mu_{t,l})^2 \|b_l^*\|^2 + \|\pi_{l+1}(v)\|^2. \end{aligned} \quad (32)$$

$$\|\pi_{g-1}(v)\|^2 = (y_{g-1} + \mu_{g,g-1})^2 \|b_{g-1}^*\|^2 + \|b_g^*\|^2 = (y_{g-1} + \mu_{g,g-1})^2 \|b_{g-1}^*\|^2 + \|\pi_g(v)\|^2. \quad (33)$$

$$\|\pi_g(v)\|^2 = \|b_g^*\|^2. \quad (34)$$

At this point, the coefficient vector of y can be sampled from y_d down to y_1 (in inverse direction) as follows:

By using Theorem 2 in [5]: $y_g = 1$ and $y_{g+1} = y_{g+2} = \dots = y_d = 0$.

By using (33) and (34): $y_{g-1} = \frac{\sqrt{\|\pi_{g-1}(v)\|^2 - \|b_g^*\|^2}}{\|b_{g-1}^*\|} - \mu_{g,g-1}$, where $\|\pi_{g-1}(v)\|^2$ is sampled as ℓ_{g-1}^2 .

By using (32) for each y_l where $1 \leq l \leq g$: $y_l = \frac{\sqrt{\|\pi_l(v)\|^2 - \|\pi_{l+1}(v)\|^2}}{\|b_l^*\|} - \sum_{t=l+1}^g y_t \mu_{t,l}$, where $\|\pi_l(v)\|^2$ is sampled in Algorithm 1 as ℓ_l^2 and also $y_{l+1}, \dots, y_g$ are sampled (found) in previous steps.

Finally, by using (14) and (13), lines 7-10 in Algorithm 1 compute the corresponding coefficients of w_t and z_{d-t+1} for $1 \leq t \leq d$.

Note: This lemma can be assumed as a well start point for entering the theorem proving methods in this context. In fact, this is possible to use formal methods by some theorem prover such as Isabelle/HOL in proving Lemma 1 by using the basic ideas, structures and propositions which are proposed in [15].

Appendix D. Proof of Lemma 3

The cutting point in full-enumeration with bounding function $\mathcal{R} = (1, 1, 1, \dots, 1)$ is always $\text{Cut} = d$. Also, by using Lemma 7 in [5] for full-enumeration, this is expected to $\mathcal{G} \approx d$. The variance of X_i is defined as follows:

$$\mathcal{V}[X_i] = E[X_i^2] - E[X_i]^2. \quad (35)$$

At first, the parameter of A_i is defined as follows: $\sum_{t=1}^d z_t^2 = \|v\|^2 \Rightarrow$

By using Theorem 2 in [5]: $\sum_{t=2}^d z_t^2 = \|v\|^2 - \|b_d^*\|^2 \Rightarrow \sum_{t=2}^d \frac{z_t^2}{\|v\|^2 - \|b_d^*\|^2} = 1 \Rightarrow$

By sampling method of (18) in this lemma (for full-enumeration): $z_{d-i+1} = (-1)^{\lfloor \text{rand}_{[0 \dots 2]} \rfloor} \sqrt{\frac{\omega_t (\|v\|^2 - \|b_d^*\|^2)}{\sum_{t=1}^{d-1} \omega_t}} \Rightarrow$

$$A_i = \sum_{t=2}^i \frac{z_t^2}{\|v\|^2 - \|b_d^*\|^2} = \frac{\sum_{t=d-i+1}^{d-1} \omega_t}{\sum_{t=1}^{d-1} \omega_t}, \quad (36)$$

where $\omega_t \leftarrow \text{Gamma}(1/2, 2)$. The value of A_i in (36) is defined as $X_i = \sum_{t=1}^i \frac{z_t^2}{R^2}$ in [1], since enumeration radius R in [1] is assumed to be $R = \|v\|$, also paper [1] dose not notice the fact by Theorem 2 in [5]. In Section 4 from [5], to show our results, we use the definition of $X_i = \sum_{t=1}^i \frac{z_t^2}{R^2}$. Therefore, the definition of $X_i = \sum_{t=1}^i \frac{z_t^2}{R^2}$ is not equal to $A_i = \sum_{t=2}^i \frac{z_t^2}{\|v\|^2 - \|b_d^*\|^2}$, especially when $\|v\| < R$.

$$\begin{aligned} X_i &= \sum_{t=1}^i \frac{z_t^2}{R^2} = \frac{\|b_d^*\|^2}{R^2} + \sum_{t=2}^i \frac{z_t^2}{R^2} = \frac{\|b_d^*\|^2}{R^2} + \frac{\|v\|^2 - \|b_d^*\|^2}{R^2} \sum_{t=2}^i \frac{z_t^2}{\|v\|^2 - \|b_d^*\|^2} \Rightarrow \\ X_i &= \frac{\|b_d^*\|^2}{R^2} + \frac{\|v\|^2 - \|b_d^*\|^2}{R^2} A_i. \end{aligned} \quad (37)$$

For invariant factors of $\frac{\|b_d^*\|^2}{R^2}$ and $\frac{\|v\|^2 - \|b_d^*\|^2}{R^2}$, by using the definition of variance and relation (37):

$$\mathcal{V}[X_i] = \mathcal{V}\left[\frac{\|b_d^*\|^2}{R^2} + \frac{\|v\|^2 - \|b_d^*\|^2}{R^2} A_i\right] = \left(\frac{\|v\|^2 - \|b_d^*\|^2}{R^2}\right)^2 \mathcal{V}[A_i]. \quad (38)$$

Also, by using the definition of variance:

$$\mathcal{V}[A_i] = E[A_i^2] - E[A_i]^2. \quad (39)$$

Step 1: In this step the value of $E[A_i^2]$ is estimated. The parameter of A_i^2 can be expanded as follows:

$$A_i^2 = \left(\sum_{t=2}^i \frac{z_t^2}{\|v\|^2 - \|b_d^*\|^2}\right)^2 = \left(\frac{\sum_{t=d-i+1}^{d-1} \omega_t}{\sum_{t=1}^{d-1} \omega_t}\right)^2, \quad \text{where } \omega_t \leftarrow \text{Gamma}(1/2, 2).$$

$$\text{For index of } x = d - i + 1: A_i^2 = \left(\frac{\sum_{t=x}^{d-1} \omega_t}{\sum_{t=1}^{d-1} \omega_t}\right)^2 = \frac{(\sum_{t=x}^{d-1} \omega_t)(\sum_{t=x}^{d-1} \omega_t)}{(\sum_{t=1}^{d-1} \omega_t)(\sum_{t=1}^{d-1} \omega_t)} \Rightarrow A_i^2 = \frac{\sum_{t=x}^{d-1} \omega_t^2 + \sum_{x \leq t \leq d-1, x \leq r \leq d-1, t \neq r} (\omega_t \omega_r)}{\sum_{t=1}^{d-1} \omega_t^2 + \sum_{1 \leq t \leq d-1, 1 \leq r \leq d-1, t \neq r} (\omega_t \omega_r)}.$$

In other side, for random variable A_i^2 , since $A_i^2 > 0$, the function $\varphi(A_i^2) = \frac{1}{A_i^2}$ is convex, therefore by using Jensen's inequality [13]: $A_i^2 > 0 \Rightarrow \frac{1}{E[A_i^2]} \leq E\left[\frac{1}{A_i^2}\right]$.

$$\begin{aligned} \frac{1}{A_i^2} &= \frac{Y_1 + Y_2}{Y_2}, \text{ where } Y_1 = \left(\sum_{t=1, r=1}^{t=x-1, r=x-1} \omega_t \omega_r\right) + \left(\sum_{t=1, r=x}^{t=x-1, r=d-1} \omega_t \omega_r\right) + \left(\sum_{t=x, r=1}^{t=d-1, r=x-1} \omega_t \omega_r\right) \text{ and} \\ Y_2 &= \left(\sum_{t=x, r=x}^{t=d-1, r=d-1} \omega_t \omega_r\right). \end{aligned}$$

The random variables of Y_1 and Y_2 are independent. By using Fact 2 in [5] and Approximation 1 in [5]:

$$\frac{1}{E[A_i^2]} \approx \frac{E[Y_1]}{E[Y_2]} + 1 \Rightarrow \frac{1}{E[A_i^2]} \approx \frac{E\left[\sum_{t=1, r=1}^{t=x-1, r=x-1} \omega_t \omega_r\right] + E\left[\sum_{t=1, r=x}^{t=x-1, r=d-1} \omega_t \omega_r\right] + E\left[\sum_{t=x, r=1}^{t=d-1, r=x-1} \omega_t \omega_r\right]}{E\left[\sum_{t=x, r=x}^{t=d-1, r=d-1} \omega_t \omega_r\right]} + 1 \Rightarrow$$

Remark 1. In approximation of variance in this context, following facts is used:

- $E[\omega_i \omega_j] = E[\omega_j] \times E[\omega_i] = 1$ for $i \neq j$ where $\omega_t \in \text{Gamma}(1/2, 2)$,
- $E[\omega_i \omega_i] = 3$ where $\omega_t \in \text{Gamma}(1/2, 2)$.

By using Remark 1: $\frac{1}{E[A_i^2]} \approx \frac{d^2-1}{i^2-1} \Rightarrow$

$$E[A_i^2] \approx \frac{i^2-1}{d^2-1}. \quad (40)$$

Step 2: In this step the value of $E[A_i]^2$ is estimated. For random variable $A_i = \frac{\sum_{t=x}^{d-1} \omega_t}{\sum_{t=1}^{d-1} \omega_t}$ where $x = d - i + 1$, since $A_i > 0$, the function $\varphi(A_i) = \frac{1}{A_i}$ is convex, therefore by using Jensen's inequality [13]: $A_i > 0 \Rightarrow \frac{1}{E[A_i]} \leq E\left[\frac{1}{A_i}\right]$.

$$\frac{1}{A_i} = \frac{\sum_{t=1}^{d-1} \omega_t}{\sum_{t=x}^{d-1} \omega_t}, \text{ where } Y_1 = \sum_{t=1}^{x-1} \omega_t \text{ and } Y_2 = \sum_{t=x}^{d-1} \omega_t.$$

The random variables of Y_1 and Y_2 are independent. By using Fact 2 in [5] and Approximation 1 in [5]:

$$\frac{1}{E[A_i]} \approx \frac{E[Y_1]}{E[Y_2]} + 1 \Rightarrow \frac{1}{E[A_i]} \approx \frac{E[\sum_{t=1}^{x-1} \omega_t]}{E[\sum_{t=x}^{d-1} \omega_t]} + 1 \Rightarrow$$

By using Remark 1: $\frac{1}{E[A_i]} \approx \frac{d-1}{i-1} \Rightarrow E[A_i] \approx \frac{i-1}{d-1} \Rightarrow$

$$E[A_i]^2 \approx \frac{(i-1)^2}{(d-1)^2} = \frac{i^2+1-2i}{d^2+1-2d}. \quad (41)$$

By using (39), (40) and (41): $\mathcal{V}[A_i] = E[A_i^2] - E[A_i]^2 = \frac{i^2-1}{d^2-1} - \frac{(i-1)^2}{(d-1)^2} = \frac{2id-2i^2-2d+2i}{d^3-d^2-d+1}$.

By using (38): $\mathcal{V}[X_i] = \left(\frac{\|v\|^2 - \|b_d^*\|^2}{R^2} \right)^2 \frac{2id-2i^2-2d+2i}{d^3-d^2-d+1}$.

References

- [1] N. Gama, P. Q. Nguyen, and O. Regev, "Lattice enumeration using extreme pruning", In Proc. EUROCRYPT '10, volume 6110 of LNCS. Springer, 2010.
- [2] Y. Chen, and P. Q. Nguyen, "BKZ 2.0: Better lattice security estimates", In International Conference on the Theory and Application of Cryptology and Information Security, pp. 1-20. Springer Berlin Heidelberg, 2011.
- [3] M. R. Albrecht, B. R. Curtis, A. Deo, A. Davidson, R. Player, E. W. Postlethwaite, F. Virdia, and T. Wunderer. "Estimate all the {LWE, NTRU} schemes!", In SCN, pages 351–357, 2018.
- [4] Shi Bai, Damien Stehlé, Weiqiang Wen, "Measuring, Simulating and Exploiting the Head Concavity Phenomenon in BKZ", ASIACRYPT 2018.
- [5] Moghissi, G., Payandeh, A. (2021). "Better Sampling Method of Enumeration Solution for BKZ-Simulation", *The ISC International Journal of Information Security*, 13(2), pp. 177-208. doi: 10.22042/iscure.2021.225886.531.
- [6] L. Ducas, "Shortest vector from lattice sieving: A few dimensions for free", In EUROCRYPT, pages 125-145, 2018.
- [7] Zhongxiang Zheng, Xiaoyun Wang, Yang Yu, "Orthogonalized lattice enumeration for solving SVP", Science China Information Sciences 2018.
- [8] Moghissi, Gholam Reza, and Ali Payandeh. "A Parallel Evolutionary Search for Shortest Vector Problem." *International Journal of Information Technology and Computer Science*, (2019).
- [9] Y. Aono, Y. Wang, T. Hayashi, and T. Takagi, "Improved progressive BKZ algorithms and their precise cost estimation by sharp simulator", In Annual International Conference on the Theory and Applications of Cryptographic Techniques, pp. 789-819. Springer, Berlin, Heidelberg, 2016.
- [10] Moghissi, Gholam Reza, and Ali Payandeh. "Using Progressive Success Probabilities for Sound-pruned Enumerations in BKZ Algorithm." *International Journal of Computer Network and Information Security* 9.9 (2018): 10.
- [11] D. Micciancio and O. Regev, "Lattice-based cryptography", In Post-quantum cryptography, pp. 147-191, Springer Berlin Heidelberg, 2009.
- [12] Moghissi, Gholam Reza, and Ali Payandeh. "Rejecting claimed speedup of $2^{\beta/2}$ in extreme pruning and revising BKZ 2.0 for better speedup", *Journal of Computing and Security*, 2021; 8(1): 65-91. doi: 10.22108/jcs.2021.121191.1044.
- [13] Jensen, Johan Ludwig William Valdemar, "Sur les fonctions convexes et les inégalités entre les valeurs moyennes", *Acta mathematica* 30 (1906): 175-193.
- [14] Gholam Reza Moghissi, Ali Payandeh, "Optimal bounding function for GNR-enumeration", *International Journal of Mathematical Sciences and Computing (IJMSC)*, Vol.8, No.1, pp. 1-17, 2022. DOI: 10.5815/ijmsc.2022.01.01.

- [15] Moghissi, Gholam Reza, and Ali Payandeh. "Formal Verification of NTRUEncrypt Scheme." *International Journal of Computer Network and Information Security* 8.4 (2016): 44.

Authors' Profiles



Gholam Reza Moghissi received the M.S. degree in the Department of ICT at Malek-e-Ashtar University of Technology, Tehran, Iran, in 2016. His researches focus on Information Security.



Ali Payandeh received the M.S. degree in Electrical Engineering from Tarbiat Modares University in 1994, and the Ph.D. degree in Electrical Engineering from K.N. Toosi University of Technology (Tehran, Iran) in 2006. He is now an assistant professor in the Department of Information and Communications Technology at the Malek-e-Ashtar University of Technology, Iran. He has published many papers in international journals and conferences. His research interests include information theory, coding theory, cryptography, security protocols, secure communications, and satellite communications.

How to cite this paper: Gholam Reza Moghissi, Ali Payandeh, "Revised Method for Sampling Coefficient Vector of GNR-enumeration Solution", *International Journal of Mathematical Sciences and Computing(IJMSC)*, Vol.8, No.3, pp. 1-20, 2022. DOI:10.5815/ijmsc.2022.03.01