Second Order Statistical Behavior of LLL and BKZ

Yang Yu[1(✉)] and Léo Ducas[2(✉)]

[1] Department of Computer Science and Technology, Tsinghua University, Beijing, China
y-y13@mails.tsinghua.edu.cn
[2] Cryptology Group, CWI, Amsterdam, The Netherlands
ducas@cwi.nl

Abstract. The LLL algorithm (from Lenstra, Lenstra and Lovász) and its generalization BKZ (from Schnorr and Euchner) are widely used in cryptanalysis, especially for lattice-based cryptography. Precisely understanding their behavior is crucial for deriving appropriate key-size for cryptographic schemes subject to lattice-reduction attacks. Current models, *e.g.* the Geometric Series Assumption and Chen-Nguyen's BKZ-simulator, have provided a decent first-order analysis of the behavior of LLL and BKZ. However, they only focused on the *average* behavior and were not perfectly accurate. In this work, we initiate a *second order analysis* of this behavior. We confirm and quantify discrepancies between models and experiments —in particular in the head and tail regions— and study their consequences. We also provide *variations* around the mean and correlations statistics, and study their impact. While mostly based on experiments, by pointing at and quantifying *unaccounted phenomena*, our study sets the ground for a theoretical and predictive understanding of LLL and BKZ performances at the second order.

Keywords: Lattice reduction · LLL · BKZ · Cryptanalysis · Statistics

1 Introduction

Lattice reduction is a powerful algorithmic tool for solving a wide range of problems ranging from integer optimization problems and problems from algebra or number theory. Lattice reduction has played a role in the cryptanalysis of cryptosystems not directly related to lattices, and is now even more relevant to quantifying the security of lattice-based cryptosystems [1,6,14].

The goal of lattice reduction is to find a basis with short and nearly orthogonal vectors. In 1982, the first polynomial time lattice reduction algorithm, LLL [15], was invented by Lenstra, Lenstra and Lovász. Then, the idea of block-wise reduction appeared and several block-wise lattice reduction algorithms [7,8,19,24] were proposed successively. Currently, BKZ is the most practical lattice reduction algorithm. Schnorr and Euchner first put forward the original BKZ algorithm in [24]. It is subject to many heuristic optimizations, such as early-abort [12], pruned enumeration [10] and progressive reduction [2,4].

C. Adams and J. Camenisch (Eds.): SAC 2017, LNCS 10719, pp. 3–22, 2018.
https://doi.org/10.1007/978-3-319-72565-9_1

All such improvements have been combined in the so-called BKZ 2.0 algorithm of Chen and Nguyen [5] (progressive strategy was improved further in later work [2]). Also, plenty of analyses [2,9,19,23,31] of BKZ algorithms have been made to explore and predict the performance of BKZ algorithms, which provide rough security estimations for lattice-based cryptography.

Despite of their popularity, the behavior of lattice reduction algorithms is still not completely understood. While there are reasonable models (e.g. the Geometric Series Assumption [25] and simulators [5]), there are few studies on the experimental statistical behavior of those algorithms, and they considered rather outdated versions of those algorithms [3,20,23]. The accuracy of the current model remains unclear.

This state of affair is quite problematic to evaluate accurately the concrete security level of lattice-based cryptosystem proposal. With the recent calls for post-quantum schemes by the NIST, this matter seems pressing.

Our Contribution. In this work, we partially address this matter, by proposing a second-order statistical (for random input bases) analysis of the behavior of reduction algorithms in practice, qualitatively and quantitatively. We figure out one more low order term in the predicted average value of several quantities such as the root Hermite factor. Also, we investigate the variation around the average behavior, a legitimate concern raised by Micciancio and Walter [19].

In more details, we experimentally study the logarithms of ratios between two adjacent Gram-Schmidt norms in LLL and BKZ-reduced basis (denoted r_i's below). We highlight three ranges for the statistical behavior of the r_i: the head ($i \leq h$), the body ($h < i < n - t$) and the tail ($i \geq n - t$). The lengths of the head and tail are essentially determined by the blocksize β. In the body range, the statistical behavior of the r_i's are similar: this does not only provide new support for the so-called Geometric Series Assumption [25] when $\beta \ll n$, but also a refinement of it applicable even when $\beta \not\ll n$. We note in particular that the impact of the head on the root Hermite factor is much stronger than the impact of the tail.

We also study the variance and the covariance between the r_i's. We observe a local correlation between the r_i's. More precisely we observe that r_i and r_{i+1} are negatively correlated, inducing a self-stabilizing behavior of those algorithms: the overall variance is less than the sum of local variances.

Then, we measure the half-volume, *i.e.* $\prod_{i=1}^{\lfloor \frac{n}{2} \rfloor} \|\mathbf{b}_i^*\|$, a quantity determining the cost of enumeration on reduced basis. By expressing the half-volume using the statistics of the r_i's, we determine that the complexity of enumeration on BKZ-reduced basis should be of the form $2^{an^2 \pm bn^{1.5}}$: the variation around average (denoted by $\pm$) can impact the speed of enumeration by a super-exponential factor.

At last, we also compare all those experimental results[1] to the simulator [5], and conclude that the simulator can predict the body of the profile and the tail

[1] The variance statistics are not comparable to the simulator [5] whose results are "deterministic", in the sense that the simulator's result starting on the Hermite Normal Form of a lattice depends only on the parameters (dimension, volume) of the lattice, and not the randomness of the lattice itself.

phenomenon qualitatively and quantitatively, but the head phenomenon is not captured. Thus it is necessary to revise the security estimation and refine the simulator.

Impact. Our work points at several inaccuracies of the current models for the behavior of LLL and BKZ, and quantifies them experimentally. It should be noted that our measured statistics are barely enough to address the question of precise prediction. Many tweaks on those algorithms are typically applied (more aggressive pruning, more subtle progressive reductions, ...) to accelerate them and that would impact those statistics. On the other hand, the optimal parametrization of heuristic tweaks is very painful to reproduce, and not even clearly determined in the literature. We therefore find it preferable to first approach stable versions of those algorithm, and minimize the space of parameters.

We would also not dare to simply guess extrapolation models for those statistics to larger blocksize: this should be the topic of a more theoretical study.

Yet, by pointing out precisely the problematic phenomena, we set the ground for revised models and simulators: our reported statistics can be used to sanity check such future models and simulators.

Source code. Our experiments heavily rely on the latest improvements of the open-source library fplll [27], catching up with the state of the art algorithm BKZ 2.0. For convenience, we used the python wrapper fpylll [28] for fplll, making our scripts reasonably concise and readable. All our scripts are open-source and available online[2], for reviewing, reproduction or extension purposes.

2 Preliminaries

We refer to [21] for a detailed introduction to lattice reduction and to [12,16] for an introduction to the behavior of LLL and BKZ.

2.1 Notations and Basic Definitions

All vectors are denoted by bold lower case letters and are to be read as row-vectors. Matrices are denoted by bold capital letters. We write a matrix $\mathbf{B}$ into $\mathbf{B} = (\mathbf{b}_1, \cdots, \mathbf{b}_n)$ where $\mathbf{b}_i$ is the i-th row vector of $\mathbf{B}$. If $\mathbf{B} \in \mathbb{R}^{n \times m}$ has full rank n, the lattice $\mathcal{L}$ generated by the basis $\mathbf{B}$ is denoted by $\mathcal{L}(\mathbf{B}) = \{\mathbf{xB} \mid \mathbf{x} \in \mathbb{Z}^n\}$. We denote by $(\mathbf{b}_1^*, \cdots, \mathbf{b}_n^*)$ the Gram-Schmidt orthogonalization of the matrix $(\mathbf{b}_1, \cdots, \mathbf{b}_n)$. For $i \in \{1, \cdots, n\}$, we define the orthogonal projection to the span of $(\mathbf{b}_1, \cdots, \mathbf{b}_{i-1})^{\perp}$ as π_i. For $1 \leq i < j \leq n$, we denote by $\mathbf{B}_{[i,j]}$ the local block $(\pi_i(\mathbf{b}_i), \cdots, \pi_i(\mathbf{b}_j))$, by $\mathcal{L}_{[i,j]}$ the lattice generated by $\mathbf{B}_{[i,j]}$.

The Euclidean norm of a vector $\mathbf{v}$ is denoted by $\|\mathbf{v}\|$. The volume of a lattice $\mathcal{L}(\mathbf{B})$ is $\mathrm{vol}(\mathcal{L}(\mathbf{B})) = \prod_i \|\mathbf{b}_i^*\|$, that is an invariant of the lattice. The first minimum of a lattice $\mathcal{L}$ is the length of a shortest non-zero vector, denoted by $\lambda_1(\mathcal{L})$. We use the shorthands $\mathrm{vol}(\mathbf{B}) = \mathrm{vol}(\mathcal{L}(\mathbf{B}))$ and $\lambda_1(\mathbf{B}) = \lambda_1(\mathcal{L}(\mathbf{B}))$.

[2] Available at https://github.com/repo-fplll/Statistical-Behavior-of-BKZ.

Given a random variable X, we denote by $\mathbf{E}(X)$ its expectation and by $\mathbf{Var}(X)$ its variance. Also we denote by $\mathbf{Cov}(X, Y)$ the covariance between two random variables X and Y. Let $\mathbf{X} = (X_1, \cdots, X_n)$ be a vector formed by random variables, its covariance matrix is defined by $\mathbf{Cov}(\mathbf{X}) = (\mathbf{Cov}(X_i, X_j))_{i,j}$.

2.2 Lattice Reduction: In Theory and in Practice

We now recall the definitions of LLL and BKZ reduction. A basis $\mathbf{B}$ is LLL-reduced with parameter $\delta \in (\frac{1}{2}, 1]$, if:

1. $|\mu_{i,j}| \leq \frac{1}{2}$, $1 \leq j < i \leq n$, where $\mu_{i,j} = \langle \mathbf{b}_i, \mathbf{b}_j^* \rangle / \langle \mathbf{b}_j^*, \mathbf{b}_j^* \rangle$ are the Gram-Schmidt orthogonalization coefficients;
2. $\delta \|\mathbf{b}_i^*\| \leq \|\mathbf{b}_{i+1}^* + \mu_{i+1,i}\mathbf{b}_i^*\|$, for $1 \leq i < n$.

A basis $\mathbf{B}$ is BKZ-reduced with parameter $\beta \geq 2$ and $\delta \in (\frac{1}{2}, 1]$, if:

1. $|\mu_{i,j}| \leq \frac{1}{2}$, $1 \leq j < i \leq n$;
2. $\delta \|\mathbf{b}_i^*\| \leq \lambda_1(\mathcal{L}_{[i, \min(i+\beta-1, n)]})$, for $1 \leq i < n$.

Note that we follow the definition of BKZ reduction from [24] which is a little different from the first notion proposed by Schnorr [26]. We also recall that, as proven in [24], LLL is equivalent to BKZ_2. Typically, LLL and BKZ are used with *Lovász parameter* $\delta = \sqrt{0.99}$ and so will we.

For high dimensional lattices, running BKZ with a large blocksize is very expensive. Heuristics improvements were developed, and combined by Chen and Nguyen [5], advertised as BKZ 2.0.[3] In this paper, we report on pure BKZ behavior to avoid perturbations due to heuristic whenever possible. Yet we switch to BKZ 2.0 to reach larger blocksizes when deemed relevant.

The two main improvements in BKZ 2.0 are called *early-abort* and *pruned enumeration*. As proven in [12], the output basis of BKZ algorithm with blocksize β would be of an enough good quality after $C \cdot \frac{n^2}{\beta^2} \left(\log n + \log \log \max \frac{\|\mathbf{b}_i^*\|}{\mathrm{vol}(\mathcal{L})^{1/n}} \right)$ tours, where C is a small constant. In our experiments of BKZ 2.0, we chose different C and observed its effect on the final basis. We also applied the pruning heuristic (see [4,10,27] for details) to speed-up enumeration, but chose a conservative success probability (95%) without re-randomization to avoid altering the quality of the output. The preprocessing-pruning strategies were optimized using the strategizer [29] of fplll/fpylll.

Given a basis $\mathbf{B}$ of an n-dimensional lattice $\mathcal{L}$, we denote by $\mathbf{rhf}(\mathbf{B})$ the root Hermite factor of $\mathbf{B}$, defined by $\mathbf{rhf}(\mathbf{B}) = \left(\frac{\|\mathbf{b}_1\|}{\mathrm{vol}(\mathcal{L})^{1/n}} \right)^{1/n}$. The root Hermite factor is a common measurement of the reducedness of a basis, e.g. [9].

Let us define the sequence $\{r_i(\mathbf{B})\}_{1 \leq i \leq n-1}$ of an n-dimensional lattice basis $\mathbf{B} = (\mathbf{b}_1, \cdots, \mathbf{b}_n)$ such that $r_i(\mathbf{B}) = \ln \left(\|\mathbf{b}_i^*\| / \|\mathbf{b}_{i+1}^*\| \right)$. The root Hermite factor $\mathbf{rhf}(\mathbf{B})$ can be expressed in terms of the $r_i(\mathbf{B})$'s:

[3] Further improvements were recently put forward [2], but are beyond the scope of this paper.

$$\mathbf{rhf}(\mathbf{B}) = \exp\left(\frac{1}{n^2}\sum_{1\leq i\leq n-1}(n-i)r_i(\mathbf{B})\right). \quad (1)$$

Intuitively, the sequence $\{r_i(\mathbf{B})\}_{1\leq i\leq n-1}$ characterizes how fast the sequence $\{\|\mathbf{b}_i^*\|\}$ decreases. Thus Eq. (1) provides an implication between the fact that the $\|\mathbf{b}_i^*\|$'s don't decrease too fast and the fact that the root Hermite factor is small. For reduced bases, the $r_i(\mathbf{B})$'s are of certain theoretical upper bounds. However, it is well known that experimentally, the $r_i(\mathbf{B})$'s tend to be much smaller than the theoretical bounds in practice.

From a practical perspective, we are more interested in the behavior of the $r_i(\mathbf{B})$'s for random lattices. The standard notion of random real lattices of given volume is based on Haar measures of classical groups. As shown in [11], the uniform distribution over integer lattices of volume V converges to the distribution of random lattices of unit volume, as V grows to infinity. In our experiments, we followed the sampling procedure of the lattice challenges [22]: its volume is a random prime of bit-length $10n$ and its Hermite normal form (see [18] for details) is sampled uniformly once its volume is determined. Also, we define a random LLL (resp. BKZ_β)-reduced basis as the basis outputted by LLL (resp. BKZ_β) applied to a random lattice given by its Hermite normal form, as described above. To speed up convergence, following a simplified progressive strategy [2,4], we run BKZ (resp. BKZ 2.0) with blocksize $\beta = 2, 4, 6, ...$ (resp. $\beta = 2, 6, 10, ...$) progressively from the Hermite normal form of a lattice.

We treat the $r_i(\mathbf{B})$'s as random variables (under the randomness of the lattice basis before reduction). For any $i \in \{1, \cdots, n-1\}$, we denote by $r_i(\beta, n)$ the random variable $r_i(\beta, n) = r_i(\mathbf{B})$, where $\mathbf{B}$ is a random BKZ_β-reduced basis, and by $\mathbb{D}_i(\beta, n)$ the distribution of $r_i(\beta, n)$. When β and n are clear from context, we simply write r_i for $r_i(\beta, n)$.

2.3 Heuristics on Lattice Reduction

Gaussian Heuristic. The Gaussian Heuristic, denoted by GAUSS, says that, for "any reasonable" subset K of the span of the lattice $\mathcal{L}$, the number of lattice points inside K is approximately $\mathrm{vol}(K)/\mathrm{vol}(\mathcal{L})$. Let the volume of n-dimensional unit ball be $V_n(1) = \frac{\pi^{n/2}}{\Gamma(n/2+1)}$. A prediction derived from GAUSS is that $\lambda_1(\mathcal{L}) \approx \mathrm{vol}(\mathcal{L})^{1/n} \cdot \mathrm{GH}(n)$ where $\mathrm{GH}(n) = V_n(1)^{-1/n}$, which is accurate for random lattices. As suggested in [10,13], GAUSS is a valuable heuristic to estimate the cost and quality of various lattice algorithms.

Random Local Block. In [5], Chen and Nguyen suggested the following modeling assumption, seemingly accurate for large enough blocksizes:

Assumption 1. [$RAND_{n,\beta}$] *Let $n, \beta \geq 2$ be integers. For a random BKZ_β-reduced basis of a random n-dimensional lattice, most local block lattices $\mathcal{L}_{[i,i+\beta-1]}$ behave like a random β-dimensional lattice where $i \in \{1, \cdots, n+1-\beta\}$.*

By $\mathrm{RAND}_{n,\beta}$ and GAUSS, one can predict the root Hermite factor of local blocks: $\mathbf{rhf}(\mathbf{B}_{[i,i+\beta-1]}) \approx \mathrm{GH}(\beta)^{\frac{1}{\beta}}$.

Geometric Series Assumption. In [25], Schnorr first proposed the Geometric Series Assumption, denoted by GSA, which says that, in typical reduced basis $\mathbf{B}$, the sequence $\{\|\mathbf{b}_i^*\|\}_{1\leq i\leq n}$ looks like a geometric series (while GAUSS provides the exact value of this geometric ratio). GSA provides a simple description of Gram-Schmidt norms and then leads to some estimations of Hermite factor and enumeration complexity [9,10]. When it comes to $\{r_i(\mathbf{B})\}_{1\leq i\leq n-1}$, GSA implies that the $r_i(\mathbf{B})$'s are supposed to be almost equal to each others. However, GSA is not so perfect, because the first and last $\mathbf{b}_i^*$'s usually violate it [3]. The behavior in the tail is well explained, and can be predicted and simulated [5].

3 Head and Tail

In [3,5], it was already claimed that for a BKZ_β-reduced basis $\mathbf{B}$, GSA doesn't hold in the first and last indices. We call this phenomenon "Head and Tail", and provide detailed experiments. Our experiments confirm that GSA holds in a strong sense in the body of the basis (*i.e.* outside of the head and tail regions). Precisely, the distributions of r_i's are similar in that region, not only their averages. We also confirm the violations of GSA in the head and the tail, quantify them, and exhibit that they are independent of the dimension n.

As a conclusion, we shall see that the head and tail have only small impacts on the root Hermite factor when $n \gg \beta$, but also that they can also be quantitatively handled when $n \not\gg \beta$. We notice that the head has in fact a stronger impact than the tail, which emphasizes the importance of finding models or simulators that capture this phenomenon, unlike the current ones that only capture the tail [5].

3.1 Experiments

We ran BKZ on many random input lattices and report on the distribution of each r_i. We first plot the average and the variance of r_i for various blocksizes β and dimensions n in Fig. 1. By superposing with proper alignment curves for the same β but various n, we notice that the head and tail behavior doesn't depend on the dimension n, but only on the relative index i (resp. $n-i$) in the head (resp. the tail). A more formal statement will be provided in Claim 1.

We also note that inside the body (*i.e.* outside both the head and the tail) the mean and the variance of r_i do not seem to depend on i, and are tempted to conclude that the distribution itself doesn't depend on i. To give further evidence of this stronger claim, we ran the Kolmogorov-Smirnov test [17] on samples of r_i and r_j for varying i, j. The results are depicted on Fig. 2, and confirm this stronger claim.

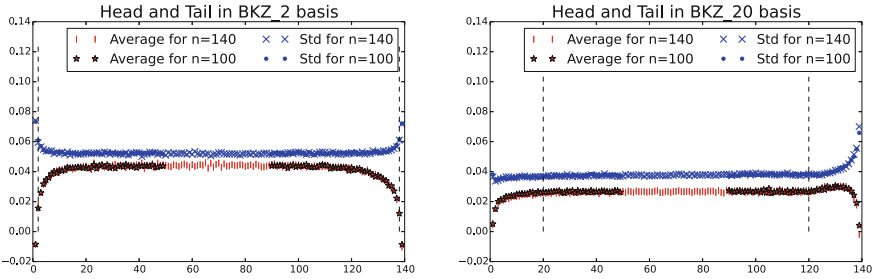

Fig. 1. Average value and standard deviation of r_i as a function of i. Experimental values measure over 5000 samples of random n-dimensional BKZ bases for $n = 100, 140$. First halves $\{r_i\}_{i \leq (n-1)/2}$ are left-aligned while last halves $\{r_i\}_{i > (n-1)/2}$ are right-aligned so to highlight heads and tails. Dashed lines mark indices β and $n - \beta$. Plots look similar in blocksize $\beta = 6, 10, 20, 30$ and in dimension $n = 80, 100, 120, 140$, which are provided in the full version.

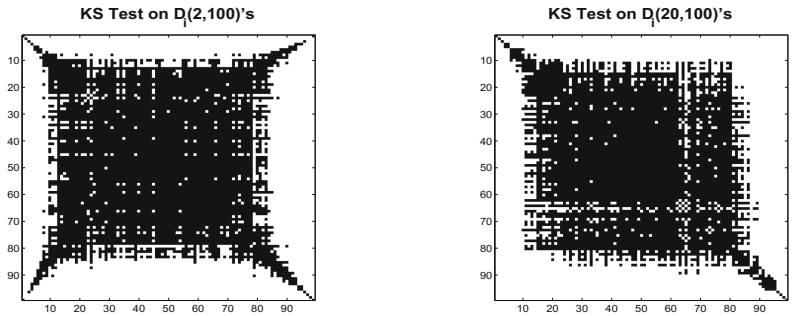

Fig. 2. Kolmogorov-Smirnov Test with significance level 0.05 on all $\mathbb{D}_i(\beta, 100)$'s calculated from 5000 samples of random 100-dimensional BKZ bases with blocksize $\beta = 2, 20$ respectively. A black pixel at position (i, j) marks the fact that the pair of distributions $\mathbb{D}_i(\beta, 100)$ and $\mathbb{D}_j(\beta, 100)$ passed Kolmogorov-Smirnov Test, *i.e.* two distributions are close. Plots in $\beta = 10, 30$ look similar to that in $\beta = 20$, which are provided in the full version.

3.2 Conclusion

From the experiments above, we allow ourselves to the following conclusion.

Experimental Claim 1. *There exist two functions $h, t : \mathbb{N} \to \mathbb{N}$, such that, for all $n, \beta \in \mathbb{N}$, and when $n \geq h(\beta) + t(\beta) + 2$:*

1. *When $i \leq h(\beta)$, $\mathbb{D}_i(\beta, n)$ depends on i and β only: $\mathbb{D}_i(\beta, n) = \mathbb{D}_i^h(\beta)$*
2. *When $h(\beta) < i < n - t(\beta)$, $\mathbb{D}_i(\beta, n)$ depends on β only: $\mathbb{D}_i(\beta, n) = \mathbb{D}^b(\beta)$*
3. *When $i \geq n - t(\beta)$, $\mathbb{D}_i(\beta, n)$ depends on $n - i$ and β only: $\mathbb{D}_i(\beta, n) = \mathbb{D}_{n-i}^t(\beta)$*

Remark 1. We only make this claim for basis that have been fully BKZ-reduced. Indeed, as we shall see later, we obtained experimental clues that this claim

would not hold when the early-abort strategy is applied. More precisely, the head and tail phenomenon is getting stronger as we apply more tours (see Fig. 4).

From now on, we may omit the index i when speaking of the distribution of r_i, implicitly implying that the only indices considered are such that $h(\beta) < i < n - t(\beta)$. The random variable r depends on blocksize β only, hence we introduce two functions of β, $e(\beta)$ and $v(\beta)$, to denote the expectation and variance of r respectively. Also, we denote by $r_i^{(h)}$(resp. $r_{n-i}^{(t)}$) the r_i inside the head (resp. tail), and by $e_i^{(h)}(\beta)$ and $v_i^{(h)}(\beta)$ (resp. $e_{n-i}^{(t)}(\beta)$ and $v_{n-i}^{(t)}(\beta)$) the expectation and variance of $r_i^{(h)}$ (resp. $r_{n-i}^{(t)}$).

We conclude by a statement on the impacts of the head and tail on the logarithmic average root Hermite factor:

Corollary 1. *For a fixed blocksize β, and as the dimension n grows, it holds that*

$$\mathbf{E}(\ln(\mathbf{rhf}(\mathbf{B}))) = \frac{1}{2}e(\beta) + \frac{d(\beta)}{n} + O\left(\frac{1}{n^2}\right), \tag{2}$$

where $d(\beta) = \sum_{i \le h} e_i^{(h)}(\beta) - \left(h + \frac{1}{2}\right) e(\beta)$.

Corollary 1 indicates that the impacts on the average root Hermite factor from the head and tail are decreasing. In particular, the tail has a very little effect $O\left(\frac{1}{n^2}\right)$ on the average root Hermite factor. The impact of the head $d(\beta)/n$, which hasn't been quantified in earlier work, is —*perhaps surprisingly*— asymptotically larger. We include the proof of Corollary 1 in Appendix A.

Below, Figs. 3 and 4 provide experimental measure of $e(\beta)$ and $d(\beta)$ from 5000 random 100-dimensional BKZ$_\beta$-reduced bases. We note that the lengths of the head and tail seem about the maximum of 15 and β. Thus we set $h(\beta) = t(\beta) = \max(15, \beta)$ simply, which affects the measure of $e(\beta)$ and $d(\beta)$ little. For the average $e(2) \approx 0.043$ we recover the experimental root Hermite factor of LLL $\mathbf{rhf}(\mathbf{B}) = \exp(0.043/2) \approx 1.022$, compatible with many other experiments [9].

To extend the curves, we also plot the experimental measure of $e(\beta)$ and $d(\beta)$[4] from 20 random 180-dimensional BKZ$_\beta$ 2.0 bases with bounded tour number $\left\lceil C \cdot \frac{n^2}{\beta^2}\left(\log n + \log\log \max \frac{\|\mathbf{b}_i^*\|}{\mathrm{vol}(\mathcal{L})^{1/n}}\right)\right\rceil$. It shows that the qualitative behavior of BKZ 2.0 is different from full-BKZ not only the quantitative one: there is a bump[5] in the curve of $e(\beta)$ when $\beta \in [22, 30]$. Considering that the success probability for the SVP enumeration was set to 95%, the only viable explanation for this phenomenon in our BKZ 2.0 experiments is the early-abort strategy: the shape of the basis is not so close to the fix-point.

[4] For BKZ 2.0, the distributions of the r_i's inside the body may not be identical, thus we just calculate the mean of those r_i's as a measure of $e(\beta)$.

[5] Yet the quality of the basis does not decrease with β in this range, as the bump on $e(\beta)$ is more than compensated by the decrease in $d(\beta)$.

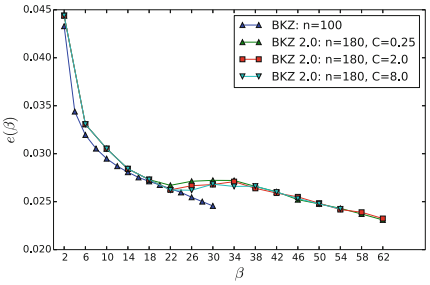

Fig. 3. Experimental measure of $e(\beta)$

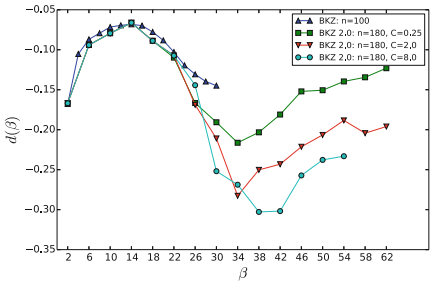

Fig. 4. Experimental measure of $d(\beta)$

4 Local Correlations and Global Variance

In the previous section, we have classified the r_i's and established a connection between the average of the root Hermite factor and the function $e(\beta)$. Now we are to report on the (co-)variance of the r_i's. Figure 5 shows the experimental measure of local variances, *i.e.* variances of the r_i's inside the body, but it is not enough to deduce the global variance, *i.e.* the variance of the root Hermite factor. We still need to understand more statistics, namely the covariances among these r_i's. Our experiments indicate that local correlations—*i.e.* correlations between r_i and r_{i+1}—are negative and other correlations seem to be zero. Moreover, we confirm the tempting hypothesis that local correlations inside the body are all equal and independent of the dimension n.

Based on these observations, we then express the variance of the logarithm of root Hermite factor for fixed β and increasing n asymptotically, and quantify the self-stability of LLL and BKZ algorithms.

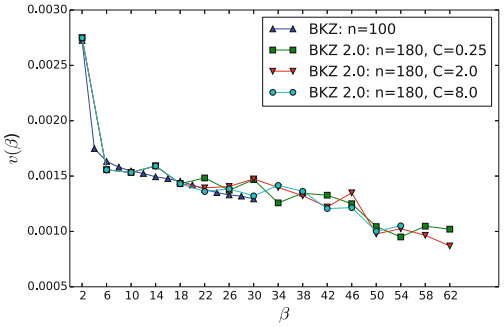

Fig. 5. Experimental measure of $v(\beta)$

4.1 Experiments

Let $\mathbf{r} = (r_1, \cdots, r_{n-1})$ be the random vector formed by random variables r_i's. We profile the covariance matrices $\mathbf{Cov}(\mathbf{r})$ for 100-dimensional lattices with BKZ reduction of different blocksizes in Fig. 6. The diagonal elements in covariance matrix correspond to the variances of the r_i's which we have studied before. Thus we set all diagonal elements to 0 to enhance contrast. We discover that the elements on the second diagonals, *i.e.* $\mathbf{Cov}(r_i, r_{i+1})$'s, are significantly negative and other elements seems very close to 0. We call the $\mathbf{Cov}(r_i, r_{i+1})$'s local covariances.

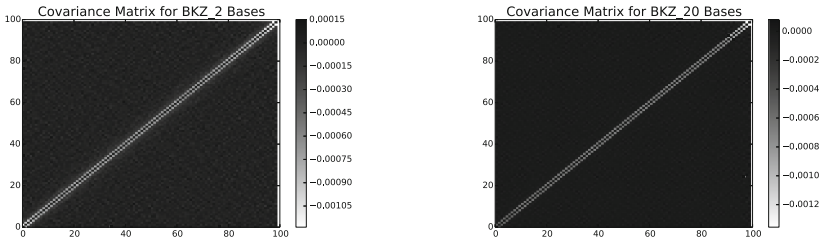

Fig. 6. Covariance matrices of $\mathbf{r}$. Experimental values measure over 5000 samples of random 100-dimensional BKZ bases with blocksize $\beta = 2, 20$. The pixel at coordinates (i, j) corresponds to the covariance between r_i and r_j. Plots in $\beta = 10, 30$ look similar to that in $\beta = 20$, which are provided in the full version.

We then plot measured local covariances in Fig. 7. Comparing these curves for various dimensions n, we notice that the head and tail parts almost coincide, and the local covariances inside the body seem to depend on β only, we will denote this value by $c(\beta)$. We also plot the curves of the $\mathbf{Cov}(r_i, r_{i+2})$'s in Fig. 7 and note that the curves for the $\mathbf{Cov}(r_i, r_{i+2})$'s are horizontal with a value about 0. For other $\mathbf{Cov}(r_i, r_{i+d})$'s with larger d, the curves virtually overlap that for the $\mathbf{Cov}(r_i, r_{i+2})$'s. For readability, larger values of d are not plotted. One thing to be noted is that the case for blocksize $\beta = 2$ is an exception. On one hand, the head and tail of the local covariances in BKZ_2 basis bend in the opposite directions, unlike for larger β. In particular, the $\mathbf{Cov}(r_i, r_{i+2})$'s in BKZ_2 basis are not so close to 0, but are nevertheless significantly smaller than the local covariances $\mathbf{Cov}(r_i, r_{i+1})$. That indicates some differences between LLL and BKZ.

Also, we calculate the average of $(n - 2\max(15, \beta))$ middle local covariances as an approximation of $c(\beta)$ for different n and plot the evolution of $c(\beta)$ in Fig. 8. The curves for different dimensions seem to coincide, which provides another evidence to support that the local covariances inside the body don't depend on n indeed. To determine the minimum of $c(\beta)$, we ran a batch of BKZ with $\beta = 2, 3, 4, 5, 6$ separately. We note that $c(\beta)$ increases with β except for $c(3) < c(2)$, which is another difference between LLL and BKZ.

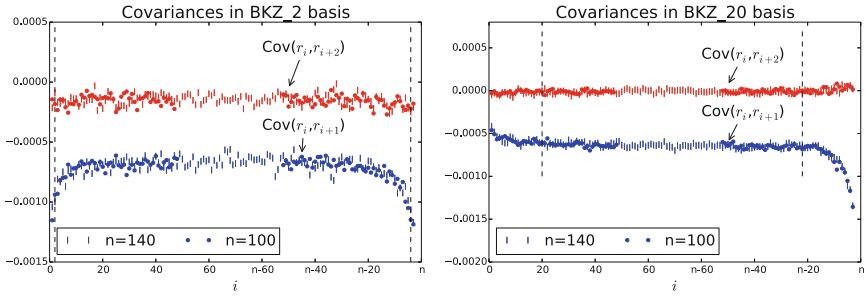

Fig. 7. $\mathbf{Cov}(r_i, r_{i+1})$ and $\mathbf{Cov}(r_i, r_{i+2})$ as a function of i. Experimental values measured over 5000 samples of random n-dimensional BKZ bases for $n = 100, 140$. The blue curves denote the $\mathbf{Cov}(r_i, r_{i+1})$'s and the red curves denote the $\mathbf{Cov}(r_i, r_{i+2})$'s. For same dimension n, the markers in two curves are identical. First halves are left aligned while last halves $\{\mathbf{Cov}(r_i, r_{i+1})\}_{i>(n-2)/2}$ and $\{\mathbf{Cov}(r_i, r_{i+2})\}_{i>(n-3)/2}$ are right aligned so to highlight heads and tails. Dashed lines mark indices β and $n-\beta-2$. Plots look similar in blocksize $\beta = 6, 10, 20, 30$ and in dimension $n = 80, 100, 120, 140$, which are provided in the full version.

Remark 2. To obtain a precise measure of covariances, we need enough samples and thus the extended experimental measure of $c(\beta)$ is not given. Nevertheless, it seems that, after certain number of tours, local covariances of BKZ 2.0 bases still tend to be negative but other covariances tend to zero.

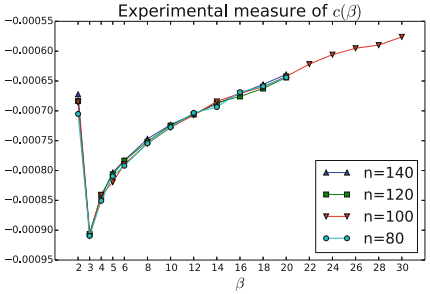

Fig. 8. Experimental measure of the evolution of $c(\beta)$ calculated from 5000 samples of random BKZ bases in different dimension n respectively.

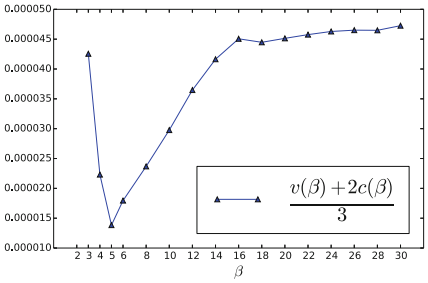

Fig. 9. Experimental measure of $\frac{v(\beta)+2c(\beta)}{3}$. The data point for $\beta = 2$, $\frac{v(2)+2c(2)}{3} \approx 0.00045$ was clipped out, being 10 times larger than all other values.

4.2 Conclusion

From above experimental observations, we now arrive at the following conclusion.

Experimental Claim 2. *Let h and t be the two functions defined in Claim 1. For all $n \in \mathbb{N}$ and $\beta > 2$ such that $n \geq h(\beta) + t(\beta) + 2$:*

1. *When* $|i-j|>1$, r_i *and* r_j *are not correlated:* $\mathbf{Cov}(r_i, r_j) = 0$
2. *When* $|i-j|=1$, r_i *and* r_j *are negatively correlated:* $\mathbf{Cov}(r_i, r_j) < 0$. *More specifically:*
   - *When* $i \leq h(\beta)$, $\mathbf{Cov}(r_i, r_{i+1})$ *depends on* i *and* β *only:* $\mathbf{Cov}(r_i, r_{i+1}) = c_i^h(\beta)$
   - *When* $h(\beta) < i < n - t(\beta)$, $\mathbf{Cov}(r_i, r_{i+1})$ *depends on* β *only:* $\mathbf{Cov}(r_i, r_{i+1}) = c(\beta)$
   - *When* $i \geq n - t(\beta)$, $\mathbf{Cov}(r_i, r_{i+1})$ *depends on* $n-i$ *and* β *only:* $\mathbf{Cov}(r_i, r_{i+1}) = c_{n-i}^t(\beta)$

One direct consequence derives from the above experimental claim is that the global variance, *i.e.* the variance of the logarithm of root Hermite factor, converges to 0 as $\Theta(1/n)$, where the hidden constant is determined by β:

Corollary 2. *For a fixed blocksize* β*, and as the dimension* n *grows, it holds that*

$$\mathbf{Var}(\ln(\mathbf{rhf}(\mathbf{B}))) = \frac{1}{3n}v(\beta) + \frac{2}{3n}c(\beta) + O\left(\frac{1}{n^2}\right). \tag{3}$$

The proof of Corollary 2 is given in Appendix B. Note that the assumption that all $\mathbf{Cov}(r_i, r_{i+d})$'s with $d > 1$ equal 0 may not be exactly true. However, the $\mathbf{Cov}(r_i, r_{i+d})$'s converge to 0 quickly as d increases, hence we may assert that the sum $\sum_{d=1}^{n-1-i} \mathbf{Cov}(r_i, r_{i+d})$ converge with n for fixed β and i inside the body. Then we still can conclude that $\mathbf{Var}(\ln(\mathbf{rhf}(\mathbf{B}))) = O(\frac{1}{n})$. The faster the $\mathbf{Cov}(r_i, r_{i+d})$'s converges to 0 as d grows, the more accurate our above approximation is. The experimental measure of $\frac{v(\beta)+2c(\beta)}{3}$ is shown in Fig. 9 and $\frac{v(\beta)+2c(\beta)}{3}$ seems to converge to a finite value $\approx 5 \times 10^{-5}$ as β grows.

5 Half Volume

We shall now study statistics on the half-volume, $\mathrm{H}(\mathbf{B}) = \prod_{i=1}^{\lfloor \frac{n}{2} \rfloor} \|\mathbf{b}_i^*\|$, of a random BKZ-reduced basis $\mathbf{B}$. As claimed in [10], the nodes in the enumeration tree at the depths around $\frac{n}{2}$ contribute the most to the total node number, for both full and regular pruned enumerations. Typically, the enumeration radius R is set to $c\sqrt{n}\cdot\mathrm{vol}(\mathbf{B})^{\frac{1}{n}}$ for some constant $c > 0$, *e.g.* $R = 1.05\cdot\mathrm{GH}(n)\cdot\mathrm{vol}(\mathbf{B})^{\frac{1}{n}}$, the number of nodes in the $\lfloor \frac{n}{2} \rfloor$ level is approximately proportional to $\frac{\mathrm{H}(\mathbf{B})}{\mathrm{vol}(\mathbf{B})^{\lceil \frac{n}{2} \rceil / n}}$, making the half-volume a good estimator for the cost of enumeration. Those formulas have to be amended in case pruning is used (see [10]), but the half-volume remains a good indicator of the cost of enumeration.

Let $\mathbf{hv}(\beta, n)$ be the random variable $\ln(\mathrm{H}(\mathbf{B})) - \frac{\lfloor \frac{n}{2} \rfloor}{n}\ln(\mathrm{vol}(\mathbf{B}))$ where $\mathbf{B}$ is a random BKZ_β-reduced basis. By the above experimental claims, we conclude the following result. The proof is shown in Appendix C.

Corollary 3 (Under previous experimental claims). *For a fixed blocksize β, let n be an integer such that $n > 2\max(h(\beta), t(\beta))$. Then, as the dimension n grows, it holds that*

$$\mathbf{E}(\mathbf{hv}(\beta, n)) = \frac{n^2}{8}e(\beta) + d'(\beta) + O\left(\frac{1}{n}\right), \tag{4}$$

where $d'(\beta) = \sum_{i \le h} \frac{i}{2}\left(e_i^{(h)}(\beta) - e(\beta)\right) + \sum_{i \le t} \frac{i}{2}\left(e_i^{(t)} - e(\beta)\right) - \frac{1}{4}\{\frac{n}{2}\}e(\beta)$, *and*

$$\mathbf{Var}(\mathbf{hv}(\beta, n)) = \frac{n^3}{48}(v(\beta) + 2c(\beta)) + O(n). \tag{5}$$

Assuming heuristically that the variation around the average of **hv** follows a Normal law, Corollary 3 implies that the complexity of enumeration on a random n-dimensional BKZ$_\beta$-reduced basis should be of the shape

$$\exp\left(n^2 x(\beta) + y(\beta) \pm n^{1.5} l \cdot z(\beta)\right) \tag{6}$$

except a fraction at most $\exp(-l^2/2)$ of random bases, where

$$x(\beta) = \frac{e(\beta)}{8}, \quad y(\beta) = d'(\beta), \quad z(\beta) = \sqrt{\frac{v(\beta) + 2c(\beta)}{48}} \tag{7}$$

and where the term $\pm n^{1.5} l \cdot z(\beta)$ accounts for variation around the average behavior. In particular, the contribution of the variation around the average remains asymptotically negligible compared to the main $\exp(\Theta(n^2))$ factor, it still introduces a super-exponential factor, that can make one particular attempt much cheaper or much more expensive in practice. It means that it could be beneficial in practice to rely partially on luck, restarting BKZ without trying enumeration when the basis is unsatisfactory.

The experimental measure of $8x(\beta)$ and $16z(\beta)^2$ has been shown in Figs. 3 and 9 respectively. We now exhibit the experimental measure of $y(\beta)$ in Fig. 10. Despite the curves for BKZ 2.0 are not smooth, it seems that $y(\beta)(=d'(\beta))$ would increase with β when β is large. However, comparing to $n^2 x(\beta)$, the impact of $y(\beta)$ on the half-volume is still much weaker.[6]

6 Performance of Simulator

In [5], Chen and Nguyen proposed a simulator to predict the behavior of BKZ. For large β, the simulator can provide a reasonable prediction of average profile, *i.e.* $\left\{\log\left(\frac{\|\mathbf{b}_i^*\|}{\mathrm{vol}(\mathcal{L})^{1/n}}\right)\right\}_{i=1}^{n}$. In this section, we will further report on the performance of the simulator qualitatively and quantitatively. Our experiments confirm that the tail still exists in the simulated result and fits the actual measure, but the head phenomenon is not captured by the simulator, affecting its accuracy for cryptanalytic prediction.

[6] The impacts of the r_i's inside the head and tail will still be significant when $\beta = O(n)$.

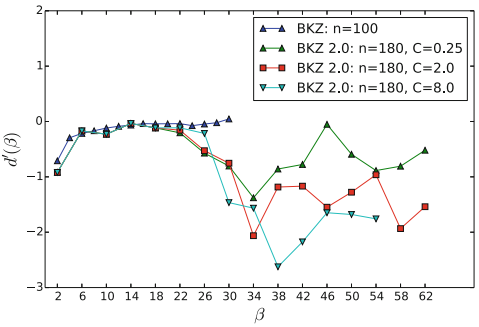

Fig. 10. Experimental measure of $y(\beta)(=d'(\beta))$

To make the simulator[7] coincide with the actual algorithm, we set the parameter $\delta = \sqrt{0.99}$ and applied a similar progressive strategy[8]. The maximum tour number corresponds to the case that $C = 0.25$ in [12], but the simulator always terminates after a much smaller number of tours.

6.1 Experiments

We ran simulator on several sequences of different dimensions and plot the average values of r_i's in Fig. 11. An apparent tail remains in the simulated result and the length of its most significant part is about β despite a slim stretch. However, there is no distinct head, which does not coincide with the actual behavior: the head shape appears after a few tours of BKZ or BKZ 2.0. Still, the r_i's inside the body share similar values, in accordance with GSA and experiments.

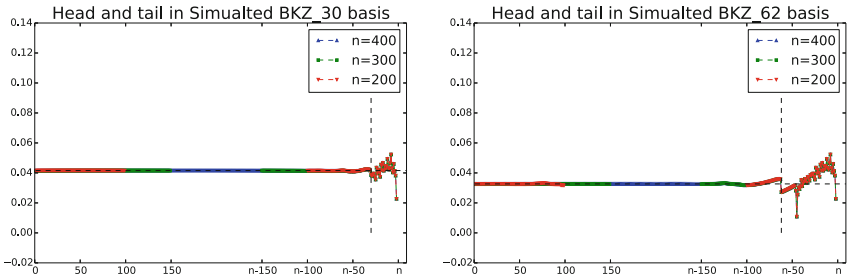

Fig. 11. Average value of r_i calculated by simulator. First halves are left aligned while last halves $\{r_i\}_{i>(n-1)/2}$ are right aligned so to highlight heads and tails. The vertical dashed line marks the index $n - \beta$ and the horizontal dashed line is used for contrast.

[7] We worked on an open-source BKZ simulator [30], with minor modifications.

[8] In simulation, the initial profile sequence is set to $(10(n-1), -10, \cdots, -10)$ and then we started from blocksize 6 and progressively ran simulator by step 2 (or 4 to simulate BKZ 2.0.). There seems to be something wrong when starting with BKZ_2.

We now compare the average experimental behavior with the simulated result. Note that the simulator is not fed with any randomness, so it does not make sense to consider variance in this comparison.

Figure 12 illustrates the comparison on $e(\beta)$. For small blocksize β, the simulator does not work well, but, as β increases, the simulated measure of $e(\beta)$ seems close to the experimental measure and both measures converge to the prediction $\ln\left(\mathrm{GH}(\beta)^{\frac{2}{\beta-1}}\right)$.

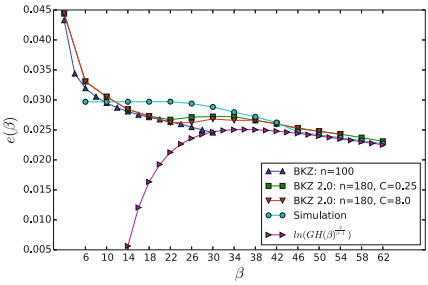

Fig. 12. Comparison on $e(\beta)$

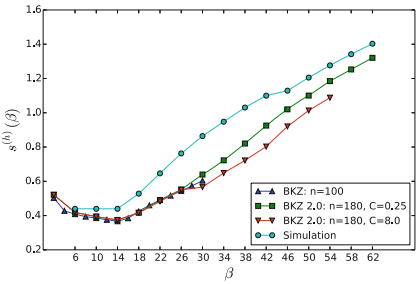

Fig. 13. Comparison on $s^{(h)}(\beta)$

Finally, we consider the two functions $d(\beta)$ and $d'(\beta)$ that are relevant to the averages of the logarithms of the root Hermite factor and the complexity of enumeration and defined in Corollarys 1 and 3 respectively. To better understand the difference, we compared the following terms $s^{(h)}(\beta) = \sum_{i \leq h} e_i^{(h)}(\beta)$, $w^{(h)}(\beta) = \sum_{i \leq h} \frac{i}{2} e_i^{(h)}(\beta)$ and $w^{(t)}(\beta) = \sum_{i \leq t} \frac{i}{2} e_i^{(t)}$ respectively, where we set $h(\beta) = t(\beta) = \max(15, \beta)$ as before. Indeed, combined with $e(\beta)$, these three terms determine $d(\beta)$ and $d'(\beta)$.

From Fig. 13, we observe that the simulated measure of $s^{(h)}(\beta)$ is greater than the experimental measure, which is caused by the lack of the head. The similar inaccuracy exists as well with respect to $w^{(h)}(\beta)$ as shown in Fig. 14. The experimental measure of $e(\beta)$ is slightly greater than the simulated measure and

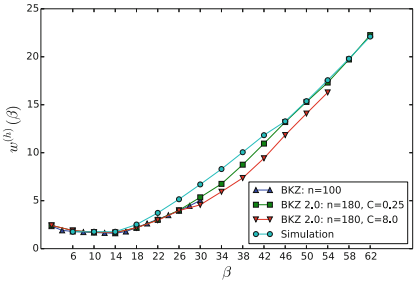

Fig. 14. Comparison on $w^{(h)}(\beta)$

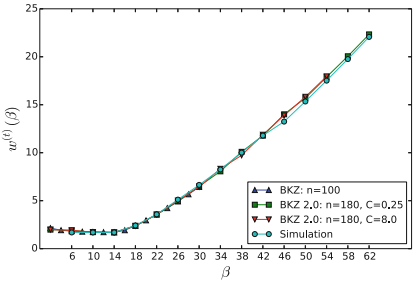

Fig. 15. Comparison on $w^{(t)}(\beta)$

thus the $e_i^{(h)}(\beta)$'s of greater weight may compensate somewhat the lack of the head. After enough tours, the head phenomenon is highlighted and yet the body shape almost remains the same so that the simulator still cannot predict $w^{(h)}(\beta)$ precisely. Figure 15 indicates that the simulator could predict $w^{(t)}(\beta)$ precisely for both large and small blocksizes and therefore the HKZ-shaped tail model is reasonable.

6.2 Conclusion

Chen and Nguyen's simulator gives an elementary profile for random BKZ$_\beta$-reduced bases with large β: both body and tail shapes are reflected well in the simulation result qualitatively and quantitatively. However, the head phenomenon is not captured by this simulator, and thus the first $\|\mathbf{b}_i^*\|$'s are not predicted accurately. In particular, the prediction of $\|\mathbf{b}_1^*\|$ that determines the Hermite factor is usually larger than the actual value, which leads to an underestimation of the quality of BKZ bases. Consequently, related security estimations need to be refined.

Understanding the main cause of the head phenomenon, modeling it and refining the simulator to include it seems an interesting and important problem, which we leave to the future work. It would also be interesting to introduce some randomness in the simulator, so to properly predict variance around the mean behavior.

Acknowledgements. We thank Phong Q. Nguyen, Jean-Christophe Deneuville and Guillaume Bonnoron for helpful discussions and comments. We also thank the SAC'17 reviewers for their useful comments. Yang Yu is supported by China's 973 Program (No. 2013CB834205), the Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDB01010600) and NSF of China (No. 61502269). Léo Ducas is supported by a Veni Innovational Research Grant from NWO under project number 639.021.645. Parts of this work were done during Yang Yu's internship at CWI.

A Proof of Corollary 1

From Eq. (1), we have:

$$\ln(\mathbf{rhf}(\mathbf{B})) = \frac{1}{n^2} \sum_{1 \le i \le n-1} (n-i) r_i(\mathbf{B}). \tag{8}$$

Taking expectations, then:

$$n^2 \mathbf{E}(\ln(\mathbf{rhf}(\mathbf{B}))) = \sum_{i \le h} (n-i) e_i^{(h)}(\beta) + \sum_{i \le t} i e_i^{(t)}(\beta) + \sum_{h < i < n-t} (n-i) e(\beta). \tag{9}$$

Note that

$$\sum_{i \le h} (n-i) e_i^{(h)}(\beta) + \sum_{i \le t} i e_i^{(t)}(\beta) = \left(\sum_{i \le t} i e_i^{(t)}(\beta) - \sum_{i \le h} i e_i^{(h)}(\beta) \right) + n \sum_{i \le h} e_i^{(h)}(\beta)$$

and

$$\sum_{h<i<n-t}(n-i)e(\beta)=\left(\frac{n^2}{2}-\frac{n(2h+1)}{2}\right)e(\beta)+\frac{(h-t)(h+t+1)}{2}e(\beta).$$

Since h and t are constant, the two terms $\left(\sum_{i\leq t} ie_i^{(t)}(\beta)-\sum_{i\leq h} ie_i^{(h)}(\beta)\right)$ and $\frac{(h-t)(h+t+1)}{2}e(\beta)$ are $O(1)$. A straightforward computation then leads to the conclusion.

B Proof of Corollary 2

We compare the variances of two sides in Eq. (8), then:

$$\begin{aligned}n^4\mathbf{Var}(\ln(\mathbf{rhf}(\mathbf{B})))&=\sum_{i=1}^{n-1}(n-i)^2\mathbf{Var}(r_i)+2\sum_{i<j}(n-i)(n-j)\mathbf{Cov}(r_i,r_j)\\&=\sum_{i=1}^{n-1}(n-i)^2\mathbf{Var}(r_i)+2\sum_{i=1}^{n-2}(n-i)(n-i-1)\mathbf{Cov}(r_i,r_{i+1}).\end{aligned}\tag{10}$$

Splitting the sum $\sum_{i=1}^{n-1}(n-i)^2\mathbf{Var}(r_i)$ into three parts, we have:

$$\sum_{i=1}^{n-1}(n-i)^2\mathbf{Var}(r_i)=\sum_{i\leq h}(n-i)^2\mathbf{Var}(r_i)+\sum_{i\geq n-t}(n-i)^2\mathbf{Var}(r_i)+\sum_{h<i<n-t}(n-i)^2\mathbf{Var}(r).\tag{11}$$

Both h and t are constant and the variances $\mathbf{Var}(r_i)$'s with $i\leq h$ or $i\geq n-t$ are also constant. Thus the two first sums are $O(n^2)$. Also, the difference $\sum_{i=1}^{n-1}(n-i)^2\mathbf{Var}(r)-\sum_{h<i<n-t}(n-i)^2\mathbf{Var}(r)$ is $O(n^2)$, then:

$$\sum_{h<i<n-t}(n-i)^2\mathbf{Var}(r_i)=\sum_{i=1}^{n-1}(n-i)^2\mathbf{Var}(r)+O(n^2)=\frac{n^3}{3}v(\beta)+O(n^2).\tag{12}$$

The sum $\sum_{i=1}^{n-2}(n-i)(n-i-1)\mathbf{Cov}(r_i,r_{i+1})$ can be split into three parts:

$$\begin{aligned}&\sum_{i\leq h}(n-i)(n-i-1)\mathbf{Cov}(r_i,r_{i+1})+\sum_{i\geq n-t}(n-i)(n-i-1)\mathbf{Cov}(r_i,r_{i+1})\\&\quad+\sum_{h<i<n-t}(n-i)(n-i-1)c(\beta).\end{aligned}\tag{13}$$

Since all $\mathbf{Cov}(r_i,r_{i+1})$'s inside the head and tail are of size $O(1)$, the first two sums are $O(n^2)$. The difference $\sum_{i=1}^{n-2}(n-i)(n-i-1)c(\beta)-\sum_{h<i<n-t}(n-i)(n-i-1)c(\beta)$ is also $O(n^2)$, then:

$$\sum_{h<i<n-t}(n-i)(n-i-1)\mathbf{Cov}(r_i,r_{i+1})=\sum_{i=1}^{n-2}(n-i)(n-i-1)c(\beta)+O(n^2)=\frac{n^3}{3}c(\beta)+O(n^2).\tag{14}$$

Combining Eq. (10), (12) and (14), we complete the proof.

C Proof of Corollary 3

Let $n' = \lfloor \frac{n}{2} \rfloor$. A routine computation leads to that:

$$\mathbf{hv}(\beta, n) = \left(1 - \frac{n'}{n}\right) \sum_{i=1}^{n'} i r_i + \frac{n'}{n} \sum_{i=n'+1}^{n-1} (n-i) r_i. \tag{15}$$

We compare the expectations of two sides in Eq. (15), then:

$$\begin{aligned} \mathbf{E}(\mathbf{hv}(\beta, n)) = &\left(1 - \frac{n'}{n}\right) \left(\sum_{i \le h} i e_i^{(h)}(\beta)\right) + \frac{n'}{n} \left(\sum_{i \le t} i e_i^{(t)}(\beta)\right) \\ &+ \left(\frac{n'(n-n')}{2} - \frac{(n-n')h(h+1) + n't(t+1)}{2n}\right) e(\beta). \end{aligned} \tag{16}$$

Since h and t are constant, the two sums $\sum_{i \le h} i e_i^{(h)}(\beta)$ and $\sum_{i \le t} i e_i^{(t)}(\beta)$ are $O(1)$. Note that $n' = \frac{n}{2} + O(1)$ and $n'(n-n') = \frac{n^2}{4} - \frac{1}{2}\{\frac{n}{2}\}$, which proves Eq. (4).

We compare the variances of two sides in Eq. (15), then:

$$\begin{aligned} \mathbf{Var}(\mathbf{hv}(\beta, n)) = &\left(1 - \frac{n'}{n}\right)^2 \left(\sum_{i \le h} i^2 v_i^{(h)}(\beta)\right) + \left(\frac{n'}{n}\right)^2 \left(\sum_{i \le t} i^2 v_i^{(t)}(\beta)\right) \\ &+ \left(\frac{n'(n-n')(2n'(n-n')+1)}{6n} - \left(1 - \frac{n'}{n}\right)^2 \sum_{i \le h} i^2 - \left(\frac{n'}{n}\right)^2 \sum_{i \le t} i^2\right) v(\beta) \\ &+ 2\left(1 - \frac{n'}{n}\right)^2 \sum_{i < n'} i(i+1)\mathbf{Cov}(r_i, r_{i+1}) \\ &+ 2\left(\frac{n'}{n}\right)^2 \sum_{i < n-n'-1} i(i+1)\mathbf{Cov}(r_{n-i}, r_{n-i-1}) \\ &+ 2\left(1 - \frac{n'}{n}\right)\left(\frac{n'}{n}\right) n'(n-n'-1)\mathbf{Cov}(r_{n'}, r_{n'+1}) \end{aligned} \tag{17}$$

We substitute all $\mathbf{Cov}(r_i, r_{i+1})$'s by $c(\beta)$, which only leads to a $O(1)$ difference. Exploiting the identity that $\sum_{i=1}^{n} i(i+1) = \frac{n(n+1)(n+2)}{3}$, we know the sum of a batch of local covariances in Eq. (17) equals $\frac{2n'(n-n')(n'(n-n')-1)}{3n} c(\beta) + O(1)$. Thus we have:

$$\mathbf{Var}(\mathbf{hv}(\beta, n)) = \frac{n'(n-n')(2n'(n-n')+1)}{6n} v(\beta) + \frac{2n'(n-n')(n'(n-n')-1)}{3n} c(\beta) + O(1), \tag{18}$$

which implies Eq. (5).

References

1. Alkim, E., Ducas, L., Pöppelmann, T., Schwabe, P.: Post-quantum key exchange—a new hope. In: USENIX Security 2016, 327–343 (2016)
2. Aono, Y., Wang, Y., Hayashi, T., Takagi, T.: Improved progressive BKZ algorithms and their precise cost estimation by sharp simulator. In: Fischlin, M., Coron, J.-S. (eds.) EUROCRYPT 2016, Part I. LNCS, vol. 9665, pp. 789–819. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49890-3_30
3. Buchmann, J., Ludwig, C.: Practical lattice basis sampling reduction. In: Hess, F., Pauli, S., Pohst, M. (eds.) ANTS 2006. LNCS, vol. 4076, pp. 222–237. Springer, Heidelberg (2006). https://doi.org/10.1007/11792086_17
4. Chen, Y.: Réduction de réseau et sécurité concrète du chiffrement complètement homomorphe. PhD thesis (2013)
5. Chen, Y., Nguyen, P.Q.: BKZ 2.0: better lattice security estimates. In: Lee, D.H., Wang, X. (eds.) ASIACRYPT 2011. LNCS, vol. 7073, pp. 1–20. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-25385-0_1
6. Ducas, L., Durmus, A., Lepoint, T., Lyubashevsky, V.: Lattice signatures and bimodal gaussians. In: Canetti, R., Garay, J.A. (eds.) CRYPTO 2013, Part I. LNCS, vol. 8042, pp. 40–56. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40041-4_3
7. Gama, N., Howgrave-Graham, N., Koy, H., Nguyen, P.Q.: Rankin's constant and blockwise lattice reduction. In: Dwork, C. (ed.) CRYPTO 2006. LNCS, vol. 4117, pp. 112–130. Springer, Heidelberg (2006). https://doi.org/10.1007/11818175_7
8. Gama, N., Nguyen, P.Q.: Finding short lattice vectors within mordell's inequality. In: STOC 2008, pp. 207–216 (2008)
9. Gama, N., Nguyen, P.Q.: Predicting lattice reduction. In: Smart, N. (ed.) EUROCRYPT 2008. LNCS, vol. 4965, pp. 31–51. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-78967-3_3
10. Gama, N., Nguyen, P.Q., Regev, O.: Lattice enumeration using extreme pruning. In: Gilbert, H. (ed.) EUROCRYPT 2010. LNCS, vol. 6110, pp. 257–278. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-13190-5_13
11. Goldstein, D., Mayer, A.: On the equidistribution of hecke points. Forum Mathematicum **15**(2), 165–189 (2003)
12. Hanrot, G., Pujol, X., Stehlé, D.: Analyzing blockwise lattice algorithms using dynamical systems. In: Rogaway, P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 447–464. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22792-9_25
13. Hanrot, G., Stehlé, D.: Improved analysis of kannan's shortest lattice vector algorithm. In: Menezes, A. (ed.) CRYPTO 2007. LNCS, vol. 4622, pp. 170–186. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74143-5_10
14. Hoffstein, J., Pipher, J., Silverman, J.H.: NTRU: a ring-based public key cryptosystem. In: Buhler, J.P. (ed.) ANTS 1998. LNCS, vol. 1423, pp. 267–288. Springer, Heidelberg (1998). https://doi.org/10.1007/BFb0054868
15. Lenstra, A.K., Lenstra, H.W., Lovász, L.: Factoring polynomials with rational coefficients. Math. Ann. **261**(4), 515–534 (1982)
16. Madritsch, M., Vallée, B.: Modelling the LLL algorithm by sandpiles. In: López-Ortiz, A. (ed.) LATIN 2010. LNCS, vol. 6034, pp. 267–281. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12200-2_25
17. Massey, F.J.: The Kolmogorov-Smirnov test for goodness of fit. J. Am. Stat. Assoc. **46**(253), 68–78 (1951)

18. Micciancio, D.: Improving lattice based cryptosystems using the hermite normal form. In: Silverman, J.H. (ed.) CaLC 2001. LNCS, vol. 2146, pp. 126–145. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-44670-2_11
19. Micciancio, D., Walter, M.: Practical, predictable lattice basis reduction. In: Fischlin, M., Coron, J.-S. (eds.) EUROCRYPT 2016, Part I. LNCS, vol. 9665, pp. 820–849. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49890-3_31
20. Nguyen, P.Q., Stehlé, D.: LLL on the average. In: Hess, F., Pauli, S., Pohst, M. (eds.) ANTS 2006. LNCS, vol. 4076, pp. 238–256. Springer, Heidelberg (2006). https://doi.org/10.1007/11792086_18
21. Nguyen, P.Q., Vallée, B.: The LLL Algorithm: Survey and applications. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-02295-1
22. Schneider, M., Gama, N.: SVP Challenge (2010). https://latticechallenge.org/svp-challenge
23. Schneider, M., Buchmann, J.A.: Extended lattice reduction experiments using the BKZ algorithm. In: Sicherheit 2010, 241–252 (2010)
24. Schnorr, C.P., Euchner, M.: Lattice basis reduction: Improved practical algorithms and solving subset sum problems. In: Budach, L. (ed.) FCT 1991. LNCS, vol. 529, pp. 68–85. Springer, Heidelberg (1991). https://doi.org/10.1007/3-540-54458-5_51
25. Schnorr, C.P.: Lattice reduction by random sampling and birthday methods. In: Alt, H., Habib, M. (eds.) STACS 2003. LNCS, vol. 2607, pp. 145–156. Springer, Heidelberg (2003). https://doi.org/10.1007/3-540-36494-3_14
26. Schnorr, C.: A hierarchy of polynomial time lattice basis reduction algorithms. Theoret. Comput. Sci. **53**(2–3), 201–224 (1987)
27. The FPLLL development team: fplll, a lattice reduction library (2016). https://github.com/fplll/fplll
28. The FPLLL development team: fpylll, a python interface for fplll (2016). Available at https://github.com/fplll/fpylll
29. The FPLLL development team: strategizer, BKZ 2.0 strategy search (2016). https://github.com/fplll/strategizer
30. Walter, M.: BKZ simulator (2014). http://cseweb.ucsd.edu/~miwalter/src/sim_bkz.sage
31. Walter, M.: Lattice point enumeration on block reduced bases. In: Lehmann, A., Wolf, S. (eds.) ICITS 2015. LNCS, vol. 9063, pp. 269–282. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-17470-9_16