Probabilistic Analysis of LLL Reduced Bases

Extended Abstract — November 23, 2009

Michael Schneider, Johannes Buchmann, and Richard Lindner

Technische Universität Darmstadt, Department of Computer Science
Hochschulstraße 10, 64289 Darmstadt, Germany
`mischnei,buchmann,rlindner@cdc.informatik.tu-darmstadt.de`

Abstract. Lattice reduction algorithms behave much better in practice than their theoretical analysis predicts, with respect to both output quality and runtime. In this paper we present a probabilistic analysis that proves an average-case bound for the length of the first basis vector of an LLL reduced bases which reflects LLL experiments much better. Additionally, we use the same method to generate average-case values for BKZ reduced bases.

Keywords: lattice reduction, LLL, worst-case bounds, average case

1 Introduction

Lattice reduction is a useful tool in cryptanalysis. Various cryptosystems are broken using lattice reduction, e.g., knapsack systems [LO85,CJL$^+$92] as well as RSA in special settings [May07]. Further on, factoring composite numbers and computing discrete logarithms is possible using lattice reduction [Sch91,May07]. Lattice reduction is also used in various cryptosystems as a normalization step during key generation, as in the Goldreich-Goldwasser-Halevi scheme [GGH97], or in reduction proofs as in [CM07]. Furthermore, the security of lattice-based cryptosystems is based on the assumed hardness of lattice basis reduction. Applications in research fields outside cryptography are optimization and operations research, i.e., solving linear integer programs [Len83]. The origin of lattice reduction is the area of number theory.

Roughly speaking, lattice reduction is the search for short vectors in a lattice. The most famous algorithm for lattice reduction is the LLL algorithm of Lenstra, Lenstra, and Lovász [LLL82]. It outputs a vector of length exponential in the lattice dimension. Practically, the most promising algorithm for LLL reduction is the L^2 algorithm by Nguyen and Stehlé [NS05]. Finding the shortest lattice vector, not just an approximation of it, is a hard problem (unless in very small dimensions). Even for polynomial approximation factors the fastest algorithm known takes time $2^{\mathcal{O}(n)}$, where n is the lattice dimension [AKS01]. In 1994, Schnorr and Euchner proposed the BKZ algorithm [SE94], which is the lattice reduction algorithm used mostly in practice today. The approximation factor achieved by BKZ is still exponential in the dimension. Theoretically, the best algorithm to find short vectors is the *slide reduction* algorithm [GN08a].

A practical comparison of lattice reduction algorithms can be found in [GN08b] and [BLR08].

One well-known fact in the area of lattice reduction is that algorithms find far shorter vectors than theoretical bounds predict. Unfortunately, the theoretical bounds are tight, i.e., there are some worst-case lattices that reach those bounds [NS05,Kan87]. This implies that it is not possible to improve the worst-case bounds. The mentioned gap shows that worst-case and average-case concerning the shortest vector's length are wide apart.

Our Contribution. In this paper we present an average-case bound for the length of the first basis vector after LLL reduction. We show a probabilistic analysis that predicts the expected length of the first basis vector of an LLL reduced basis to be around $1.0193^{n-1} \cdot \det(L)^{1/n}$, whereas the worst-case analysis only yields the bound $1.0782^{n-1} \cdot \det(L)^{1/n}$. Here n is the lattice dimension and $\det(L)$ is a lattice constant. We show that the average-case bound is close to what algorithms reach in practice. To obtain our result, we assume that some characteristic numbers that arise during lattice reduction behave like random variables. Those numbers represent the degree of reducedness. Among others, the Gram-Schmidt coefficients are used as indicators for the reducedness. Assuming those coefficients to be random variables gives rise to a probabilistic analysis of the LLL reduced bases. For this analysis we present probability distributions and claim that the characteristic numbers follow this distributions. The distributions are deduced from experiments that we performed on random lattices chosen as in [GN08b,NS06] and on modular lattices like those in [BLR08]. In the second part of the paper we generate comparable average-case estimates for BKZ reduced bases. Our new bounds reflect the practical results of lattice reduction far better than the existing worst-case bounds and will be helpful in estimating key sizes of lattice based cryptosystems.

Related Work. First approaches concerning the gap between theory and practice of lattice reduction were made in [NS06] and [GN08b]. Both papers analyse the practical behaviour of reduction algorithms by evaluating their experiments. The authors of [GN08b] assume that, when starting with a random lattice following the probability distribution of [GM03], the output length of the first basis vector is around $1.02^n \cdot \det(L)^{1/n}$. In this paper we assume a different probability distribution of an LLL reduced basis *after* LLL reduction. This distribution is then used to theoretically derive a prediction of the first lattice vector in the average case. In [VV07,DFV97] Valleé *et al.* present an average-case analysis of the Gaussian algorithm for lattice reduction only in dimension 2.

Although it sounds similar, we do not address the famous worst-case to average-case reduction for lattices [Ajt96]. Our lattices are already reduced, the dimension of average-case and worst-case lattices is the same, we do not care about hardness assumptions of lattice reduction.

2 Preliminaries

Let $n, d \in \mathbb{N}$, $n \leq d$, $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{R}^d$ linearly independent. Then $L(\mathbf{B}) = \{\sum_{i=1}^n x_i \mathbf{b}_i : x_i \in \mathbb{Z}\}$ is the lattice spanned by $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$. $L(\mathbf{B})$ has dimension n, $\mathbf{B}$ is a basis of the lattice. Such a basis is uniquely determined up to unimodular transformations. We write L instead of $L(\mathbf{B})$ if it is clear which basis is concerned. The first successive minimum $\lambda_1(L)$ is the length of the shortest vector of a lattice. The lattice determinant $\det(L(\mathbf{B}))$ is defined as $\sqrt{\det(\mathbf{B}\mathbf{B}^t)}$. It is invariant under basis changes. For full-dimensional lattices ($n = d$) there is $\det(L(\mathbf{B})) = |\det(\mathbf{B})|$ for every basis $\mathbf{B}$. When writing $\|\cdot\|$ we refer to the usual Euclidean norm.

Denote the Gram-Schmidt-orthogonalization (GSO) with $\mathbf{b}_i^* = \pi_i(\mathbf{b}_i)$ where $\pi_i(\mathbf{b}) \to \mathrm{span}(\mathbf{b}_1 \ldots \mathbf{b}_{i-1})^\perp$ is the orthogonal projection. The GSO is calculated via $\mathbf{b}_i^* = \mathbf{b}_i - \sum_{j=1}^{i-1} \mu_{i,j} \mathbf{b}_j^*$ where $\mu_{i,j} = \mathbf{b}_i^T \mathbf{b}_j^* / \|\mathbf{b}_j^*\|^2$ for all $1 \leq j \leq i \leq n$. We know that $\prod_{k=1}^n \|\mathbf{b}_k^*\| = \det(L(\mathbf{B}))$.

Lattice Reduction. Creating a basis consisting of short and nearly orthogonal vectors is the goal of lattice reduction. A more detailed notion of a reduced lattice is the following. A basis $[\mathbf{b}_1, \ldots, \mathbf{b}_n]$ is called δ-LLL reduced with $\delta \in (\frac{1}{4}, 1]$, if

$$|\mu_{i,j}| \leq 0.5 \qquad \text{for } 1 \leq j < i \leq n \text{ and}$$
$$\delta \|\mathbf{b}_{i-1}^*\|^2 \leq \|\mathbf{b}_i^*\|^2 + \mu_{i,i-1}^2 \|\mathbf{b}_{i-1}^*\|^2 \qquad \text{for } i = 2, \ldots, n.$$

A basis satisfying the first condition is called *size-reduced*. The second condition is called the Lovász-condition. Notice that the LLL definition is parameterized by δ. Throughout this paper, we are using $\delta = 0.99$, which is a common choice.

Hard Lattice Problems and Algorithms. There are several problems on lattices that are supposed to be or proven to be hard [MG02]. The most famous problem is the shortest vector problem (SVP). The goal of γ-SVP is to find an (approximate) shortest non-zero vector in the lattice, namely a vector $\mathbf{0} \neq \mathbf{v} \in L$ with $\|\mathbf{v}\| \leq \gamma \lambda_1(L)$, where $\gamma \geq 1$ is the approximation factor. It is possible to formulate the problem in every norm, the most usual norm is the Euclidean norm, that we are using throughout this paper.

As the length of the shortest vector $\lambda_1(L)$ might not be known, it might be hard to control the approximation factor of SVP. Therefore it is common practice to use the Hermite-SVP variant: given a $\gamma \geq 1$, find a non-zero vector $\mathbf{v} \in L$ with $\|\mathbf{v}\| \leq \gamma \cdot \det(L)^{1/n}$. Having reduced a basis $\mathbf{B}$ one can easily calculate the reached Hermite factor using $\gamma_{\mathrm{Hermite}} = \|\mathbf{b}_1\| / \det(L)^{1/n}$.

The approximate SVP as well as the Hermite SVP was solved by Lenstra, Lenstra and Lovász in [LLL82] for factors γ exponential in the lattice dimension n. Their LLL algorithm requires time $\mathcal{O}(d^5 n \log^3 B)$ using floating point arithmetic with precision $\mathcal{O}(d \log B)$, where B is an upper bound for the length of the input vectors, and outputs a basis whose first vector approximates the shortest lattice vector with an approximation factor exponential in the lattice dimension. More concretely, it can be proved that $\|\mathbf{b}_1\| \leq (\delta - 1/4)^{(1-n)/4} \cdot \det(L)^{1/n}$

[LLL82] after LLL reduction, where δ is the LLL parameter. In other words, LLL provably reaches a Hermite factor of $(\delta - 1/4)^{(n-1)/4}$. For $\delta = 0.99$ this upper bound is $\gamma_{\mathsf{LLL}} = 1.0782^{n-1}$. There exist worst-case bases that reach the upper bound, therefore the LLL worst-case bound is tight.

In [NS05] the authors propose a new algorithm for LLL reduction, called L^2 algorithm. It also outputs δ-LLL reduced bases, and it runs in time $\mathcal{O}(d^4 n(d + \log B) \log B)$. The main advantage of this algorithm is the lower precision of $\mathcal{O}(d \log_2 3)$ that is required for termination. The L^2 algorithm is implemented in the fpLLL library [Ste]. In [SE94] the authors introduce the idea of deep inserting the size-reduced vector into the basis, called deepLLL. This algorithm variant finds shorter vectors, at the expense of an increase in runtime. The same paper presents the BKZ algorithm, that is a blockwise variant of the LLL algorithm. BKZ is today's best algorithm for lattice reduction in practice. It is parameterized with block size parameter β and provably reaches a lattice vector with length $\|\mathbf{b}_1\| \leq (\gamma_\beta)^{\frac{n-1}{\beta-1}} \cdot \lambda_1(L)$ [Sch94], where γ_β is the Hermite constant in dimension β (do not confound with the Hermite *factor* that we are using throughout this paper). Higher block sizes lead to shorter vectors, at the expense of an increased runtime. Every BKZ reduced basis is always also LLL reduced [SE94].

Practical Behaviour. In practice however, lattice reduction algorithms behave much better than theory would suppose: in the average case they find much shorter vectors than theoretical worst-case bounds suggest. In [GN08b] Gama and Nguyen give a practical analysis of LLL and BKZ using the established implementation of Shoup's NTL library [Sho]. The authors state that a Hermite factor of 1.01^n and an approximation factor of 1.02^n in high lattice dimension (e.g. dimension 500) is within reach today (using BKZ or deepLLL) but a Hermite factor of 1.005^n in dimension around 500 is totally out of reach. The Hermite factor reached by LLL experimentally is $\gamma_{\exp} \approx 1.02^n$. BKZ algorithms reach a Hermite factor of $\gamma_{\mathsf{BKZ}} \approx 1.01^n$. The deep insertion variant of LLL was observed in [BW02,NS06] and [GN08b]. The average Hermite factor reached by deepLLL is observed to be 1.012^n (maximal insertion depth not given) in [NS06] and 1.011^n with maximal insertion depth 50 in [GN08b], respectively.

3   LLL on the Average Revisited

We have seen that the output result of an LLL algorithm is much smaller than the worst-case bound predicts. In dimension 500 the predicted length of $\mathbf{b}_1$ in a δ-LLL reduced basis with $\delta = 0.99$ is $1.0782^{499} \approx 2 \cdot 10^{16} \cdot \det(L)^{1/n}$, whereas practically the resulting norm would be around $1.02^{500} \approx 2 \cdot 10^4 \cdot \det(L)^{1/n}$. In this section we analyze the worst-case bound, show why it is far away from practice, and present a better average-case bound for the length of $\mathbf{b}_1$ after LLL reduction.

Both LLL conditions introduce a possible error in the worst-case analysis. The size-reduction condition as well as the Lovász condition can be more or less violated, but will seldom be satisfied with equality. We introduce two different

kinds of random variables that each indicate one of the two LLL conditions. Using these random variables we will then compute an expectation value for the norm of the first basis vector in an LLL reduced basis that is very close to the actual behaviour of LLL algorithms in practice.

3.1 Probabilistic Size Reduction

Remember that in an LLL reduced basis it is required that $|\mu_{i,j}| \leq 0.5$ for $0 \leq j < i \leq n$. For the theoretical LLL analysis one assumes [LLL82] that all $\mu_{i,i-1}$ match the worst case, i.e. $|\mu_{i,i-1}| = 0.5$. Our new idea is to replace this assumption by a more realistic, probabilistic distribution of the Gram-Schmidt coefficients. The challenge is to find a suitable distribution function of those random variables representing the Gram-Schmidt coefficients. In [NS06] the authors state that the coefficients $\mu_{i,i-1}$ are not uniformly distributed in the area $[-0.5, 0.5]$. There, the authors also present some results of how the μ's are distributed.

In order to get a better impression of the distribution of the $\mu_{i,i-1}$ we performed some experiments on random lattices like those used in [NS06] and [GN08b] as well as on the modular lattices of [BLR08]. We created 200 random lattices in dimension 100 and additionally used the challenge lattices in dimension 200 - 475[1]. For all experiments we used the reduction parameter $\delta = 0.99$. We discovered that in other dimensions, the results are the same as in dimension 100, therefore we fixed the dimension at a dimension where bigger numbers of lattices can be reduced in a reasonable amount of time.

It turns out that the values $\mu_{i,i-1}$ are distributed along a polynomial of degree four, namely $p(y) = 53.85692y^4 + 1.57202y^2 + 0.19579$ in the range $[-0.5, 0.5]$. We expect the values $\mu_{i,i-1}$ to be represented by random variables that we denote by Y. Figure 1 shows the experimental data and the fitting polynomial $p(y)$. The polynomial was created using the least-squares method; $p(x)$ minimizes the term $\sum_y (p(y) - y_i)^2$ over all symmetric polynomials of degree four. Here, y_i represents the experimental data points. The polynomial $p(y)$ is created such that $\int_{-\infty}^{\infty} p(y)\, dy = 1$, which allows us to use $p(y)$ as density function of a probability distribution:

$$p(y) = \begin{cases} 53.85692y^4 + 1.57202y^2 + 0.19579 & \text{if } y \in [-0.5, 0.5] \\ 0 & \text{otherwise} \end{cases}. \quad (1)$$

We use the random variables Y to derive an average-case estimate for the norm of the first basis vector after LLL reduction. The result is presented in the first part of Theorem 1.

3.2 Probabilistic Lovász Condition

Secondly, we want to get an idea of how strong the Lovász condition is violated. For this we LLL reduce our random lattice bases, then compute the

---

[1] www.latticechallenge.org
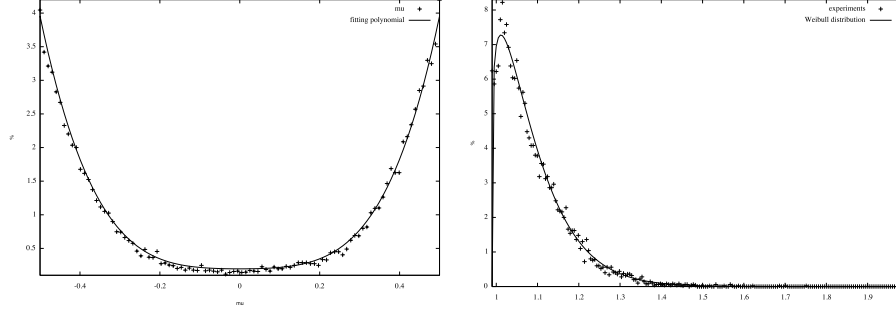
Fig. 1. Distribution of the values $\mu_{i,i-1}$ and the fitting polynomial $p(y)$ of (1).

Fig. 2. Distribution of the values x_i and the fitting curve $f(x)$ of (2).

Gram-Schmidt orthogonalization and compute

$$x_i = \frac{\left\|\pi_{i-1}(\mathbf{b}_i)\right\|^2}{\left\|\pi_{i-1}(\mathbf{b}_{i-1})\right\|^2} = \frac{\left\|\mathbf{b}_i^* + \mu_{i,i-1}\mathbf{b}_{i-1}^*\right\|^2}{\left\|\mathbf{b}_{i-1}^*\right\|^2} .$$

Remember that the Lovász condition is $\delta \left\|\mathbf{b}_{i-1}^*\right\|^2 \leq \left\|\mathbf{b}_i^* + \mu_{i,i-1}\mathbf{b}_{i-1}^*\right\|^2$. We know that $\delta \leq x_i$ because the basis is LLL-reduced. For the worst-case analysis one assumes only that $x_i \geq 0.99$. We will now present a more detailed analysis of how the x_i behave in practice.

In order to describe this attitude of an LLL reduced basis we introduce some second random variables X describing the values x_i of our experiments. We choose a Weibull density function for those variables that fit the experimental data. The density of a Weibull distribution is given by

$$f(x) = \begin{cases} \alpha \cdot \beta \cdot (x-c)^{\beta-1} \cdot e^{-\alpha(x-c)^\beta} & \text{for } x \geq c \\ 0 & \text{otherwise} \end{cases} . \tag{2}$$

Using Maximum-likelihood fitting we derive parameters $\alpha = 14.3752$, $\beta = 1.18$, and shift in x-direction with $c = 0.99$. Figure 2 shows the density function fitting the experimental values. The second part of Theorem 1 shows the impact of choosing X as random variables.

3.3 Probabilistic Analysis of LLL Reduced Bases

Next we suppose that both the x_i and the $\mu_{i,i-1}$ can be represented by independent random variables X and Y with density function given by $f(x)$ and $p(y)$ (Formulae (2) and (1)). We construct a two-dimensional random variable (X,Y) that has density function $d(x,y) = f(x) \cdot p(y)$. This implies the assumption that X and Y are independent random variables.

Our main result is presented as Theorem 1. There we present the expectation of $\|\mathbf{b}_1\|$ when choosing either μ, X, or both μ and X randomly. Remember that

the LLL worst-case bound is $1.0782^{n-1} \cdot \det(L)^{1/n}$ and practical LLL algorithms output vectors with length $1.02^n \cdot \det(L)^{1/n}$. The basis of the proof is the same as in the original LLL proof [LLL82].

Theorem 1. *Suppose that a basis $[\mathbf{b}_1 \dots \mathbf{b}_n]$ is chosen arbitrarily and an* LLL *algorithm is performed on the basis.*

a) *Suppose that after* LLL *reduction, coefficients $\mu_{i,i-1}$ are represented by independent random variables Y and their probability distribution is given by $p(y)$. Then the expectation of the norm of the first lattice vector after* LLL *reduction is*

$$E(\|\mathbf{b}_1\|) \lessapprox 1.0474^{n-1} \cdot \det(L)^{1/n} \,. \tag{3}$$

b) *Suppose that after* LLL *reduction, quotients x_i are represented by independent random variables X and their probability distribution is given by $f(x)$. Then the expectation of the norm of the first lattice vector after* LLL *reduction is*

$$E(\|\mathbf{b}_1\|) \lessapprox 1.0462^{n-1} \cdot \det(L)^{1/n} \,. \tag{4}$$

c) *Suppose that after* LLL *reduction, quotients x_i and coefficients $\mu_{i,i-1}$ are represented by independent random variables (X, Y) and their probability distribution is given by $f(x) \cdot p(y)$. Then the expectation of the norm of the first lattice vector after* LLL *reduction is*

$$E(\|\mathbf{b}_1\|) \approx 1.0193^{n-1} \cdot \det(L)^{1/n} \,. \tag{5}$$

Proof. Remember that $0 \leq \mu_{i,i-1}^2 \leq 1/4$ and $x_i \geq 0.99$. We start with an LLL reduced basis $\mathbf{B}$:

$$\delta \left\|\mathbf{b}_{i-1}^*\right\|^2 \leq \mu_{i,i-1}^2 \left\|\mathbf{b}_{i-1}^*\right\|^2 + \|\mathbf{b}_i^*\|^2 \qquad \forall\, i = 2, \dots, n$$

$$\left\|\mathbf{b}_{i-1}^*\right\|^2 \leq (\delta - \mu_{i,i-1}^2)^{-1} \|\mathbf{b}_i^*\|^2 \qquad \forall\, i = 2, \dots, n$$

$$\|\mathbf{b}_1^*\|^2 \leq \prod_{i=1}^{k-1} (\delta - \mu_{i,i-1}^2)^{-1} \|\mathbf{b}_k^*\|^2 \qquad \forall\, k = 1, \dots, n$$

Multiplying both sides for all k from 1 to n gives[2]

$$\|\mathbf{b}_1^*\|^{2n} \leq \prod_{k=1}^{n} \Big(\prod_{i=1}^{k-1} (\delta - \mu_{i,i-1}^2)^{-1} \Big) \cdot \det(L)^2 \,.$$

Applying the logarithm on both sides we get

$$\ln(\|\mathbf{b}_1^*\|) \leq -\frac{1}{2n} \sum_{k=1}^{n} \sum_{i=1}^{k-1} \ln(\delta - \mu_{i,i-1}^2) + \frac{1}{n} \ln(\det(L)) \,.$$

Now we calculate the expectation values of $\ln(\|\mathbf{b}_1^*\|)$. Therefore we consider the three different cases:

---

[2] Remember that $\prod_{i=1}^{n} \|\mathbf{b}_i^*\| = \det(L)$.

a) We suppose that the $\mu_{i,i-1}$ behave like random variables Y, with distribution given by the polynomial $p(y)$. Then we get

$$E(\ln(\|\mathbf{b}_1^*\|)) \leq -\frac{1}{2n} \sum_{k=1}^{n} \sum_{i=1}^{k-1} E(\ln(\delta - Y^2)) + \frac{1}{n} \ln(\det(L)).$$

It is $E(\ln(\delta - Y^2)) = \int_{-\infty}^{\infty} \ln(\delta - y^2) \cdot p(y)\, dy$. We use the software package GNU Octave [Oct] for the calculation using Gaussian quadrature and compute that $E(\ln(\delta - Y^2)) = -0.18519$ for $\delta = 0.99$. Using this we get

$$E(\ln(\|\mathbf{b}_1^*\|)) \leq -\frac{1}{2n} \sum_{k=1}^{n} \sum_{i=1}^{k-1} -0.18519 + \frac{1}{n} \ln(\det(L))$$

$$= \frac{0.18519 \cdot n(n-1)}{2 \cdot 2n} + \frac{1}{n} \ln(\det(L))$$

$$= 0.046297 \cdot (n-1) + \frac{1}{n} \ln(\det(L)).$$

b) For the second case, where the x_i are considered to be random X, the Lovász condtion is replaced by $x_i \left\|\mathbf{b}_{i-1}^*\right\|^2 = \mu_{i,i-1}^2 \left\|\mathbf{b}_{i-1}^*\right\|^2 + \left\|\mathbf{b}_i^*\right\|^2$. So we get equality up to the step where we replace $\mu_{i,i-1}^2$ by 0.25. Finally to compute the expectation value we use again Gaussian quadrature. The result is

$$E(\ln(X - 0.25)) = \int_{0.99}^{\infty} \ln(x - 0.25) \cdot f(x)\, dx = -0.18051.$$

Analogue to case a) this leads to

$$E(\ln(\|\mathbf{b}_1^*\|)) \leq 0.045128 \cdot (n-1) + \frac{1}{n} \ln(\det(L)).$$

c) When assuming that both indicators behave like random variables (X, Y), we get equality in all inequations:

$$E(\ln(\|\mathbf{b}_1^*\|)) = -\frac{1}{2n} \sum_{k=1}^{n} \sum_{i=1}^{k-1} E(\ln(X - Y^2)) + \frac{1}{n} \ln(\det(L)).$$

Now we are using two-dimensional Gaussian quadrature of GNU Octave (the method *quad2dg*) to compute the integral.

$$E(\ln(X - Y^2)) = \int_{0.99}^{\infty} \int_{-0.5}^{0.5} \ln(x - y^2) \cdot f(x)p(y)\, dy\, dx = -0.076560$$

$$\Rightarrow E(\ln(\|\mathbf{b}_1^*\|)) = 0.019140 \cdot (n-1) + \frac{1}{n} \ln(\det(L)).$$

Finally in each of the three cases we compute $E(\|\mathbf{b}_1\|) = E(\|\mathbf{b}_1^*\|) \approx \exp(E(\ln(\|\mathbf{b}_1^*\|)))$:

$$E(\|\mathbf{b}_1\|) \text{ a) } \lessapprox 1.0474^{n-1} \cdot \det(L)^{1/n}$$

$$\text{b) } \lessapprox 1.0462^{n-1} \cdot \det(L)^{1/n}$$

$$\text{c) } \approx 1.0193^{n-1} \cdot \det(L)^{1/n}$$

$\square$

Notice that this results directly show the corresponding Hermite factors. The average-case Hermite factors are $\gamma_{a)} = 1.0474^{n-1}$, $\gamma_{b)} = 1.0462^{n-1}$, and $\gamma_{c)} = 1.0193^{n-1}$, respectively. The error introduced by Gaussian quadrature is negligible.

3.4 Interpretation of the Results

The original LLL worst-case Hermite factor for $\delta = 0.99$ is

$$\|\mathbf{b}_1\| / \det(L)^{1/n} \leq (0.99 - 1/4)^{(1-n)/4} \approx 1.0782^{n-1} . \qquad (6)$$

In [NS06] Nguyen and Stehlé show experimentally that practical LLL algorithms reach an average value of

$$\|\mathbf{b}_1\| / \det(L)^{1/n} \approx 1.02^n . \qquad (7)$$

The experiments of Gama and Nguyen [GN08b] show the same behaviour of LLL algorithms. Figure 3 gives an illustration of all Hermite factors, i.e. the Figure shows the norm of the first lattice vector divided out the common $\det(L)^{1/n}$-part.
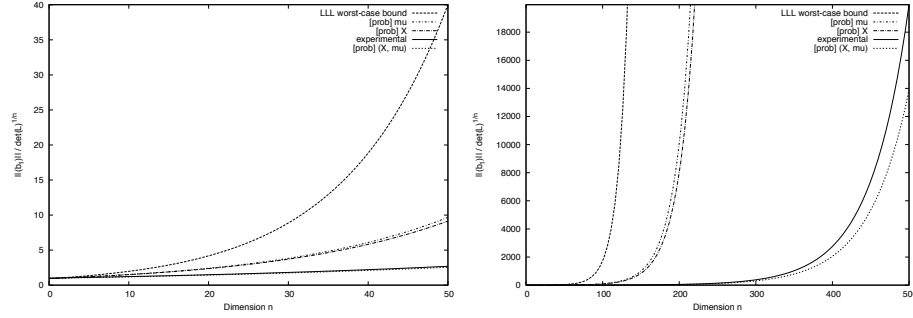
Fig. 3. Length of the first basis vector after LLL reduction; LLL worst-case bound (6), experimental value (7) of [NS06] and new average-case bounds (3)(4)(5). The $\det(L)^{1/n}$-part is omitted.

Clearly the LLL worst-case bound is not a good candidate to predict the length of the first lattice vector. One can see that the random X and Y have similar impact on the prediction. The results (1.0474^{n-1} and 1.0462^{n-1}) are close together. But still the experimental data are far away from the prediction if we only use one of both random variables. If we use both at the same time, the prediction gets quite good. Indeed, the expectation of the norm of the first basis vector gets a bit too short in that case. This might be caused by our assumption that all random variables are independent from each other. The density function of the combined, two-dimensional random variable (X, Y) might differ from the assumed $f(x) \cdot p(y)$. But still our theoretical prediction of the norm of the first basis vector is very close to what practical algorithms output.

4 Probabilistic **BKZ** Analysis

In this section we apply our probabilistic analysis to β-BKZ reduced bases. For those bases, the worst-case bounds are hard to compare to experimental results. The known bound for the first basis vector in a β-BKZ reduced basis is

$$\|\mathbf{b}_1\| \leq \gamma_\beta^{(n-1)/(\beta-1)} \cdot \lambda_1(L) \,,$$

where γ_β is the Hermite constant in dimension β which is known for values $\beta \leq 8$ [Cas71]. It is known that $\gamma_\beta \leq \frac{2}{3}\beta$ for $\beta \geq 2$ and that $\gamma_\beta \leq \frac{0.872\beta}{\pi e}(1 + o(1))$ for $\beta \to \infty$ [KL78]. With $\lambda_1 \leq \sqrt{n}\det(L)^{1/n}$ we can calculate a worst-case bound $\|\mathbf{b}_1\| \leq (\frac{2}{3}\beta)^{(n-1)/(\beta-1)}\sqrt{n}\det(L)^{1/n}$, which for $\beta = 20$ becomes $\|\mathbf{b}_1\| \leq 1.1461^{n-1}\sqrt{n}\det(L)^{1/n}$. This bound is much worse than the LLL worst-case bound, therefore it is not useful for our analysis. To our knowledge, there is no theoretical bound on the Hermite factor reached by BKZ.

Practically, [GN08b] shows that $\|\mathbf{b}_1\| \approx 1.01^n \cdot \det(L)^{1/n}$ after BKZ-reduction, where the block size β is not explicitly given. We use this for comparison in our following probabilistic analysis.

4.1 Probability Distributions

We use the same lattices as in the previous chapter and reduce them using BKZ-20 of NTL (BKZ with blocksize 20). We use the same fitting methods as before, namely least-squares method for the $\mu_{i,i-1}$. For the density function of the Gram-Schmidt coefficients we get

$$p_2(y) = \begin{cases} 12.59299y^4 + 8.89436y^2 + 0.10139 & \text{if } y \in [-0.5, 0.5] \\ 0 & \text{otherwise} \end{cases}, \quad (8)$$

and for the x_i's the Maximum Likelihood fitting gives us Weibull parameters $\alpha = 24.7316$, $\beta = 1.5323$, and shift in x-direction with $c = 0.99$. Denote $f_2(x)$ the Weibull distribution function (2) using those new parameters. We use this distribution functions to derive average-case values for the norm of the first basis vector after BKZ reduction.

4.2 Probabilistic Analysis

We advance exactly the same as in the case of Theorem 1. We suppose that either the $\mu_{i,i-1}$, the x_i, or both of them behave like random variables following the distributions $p_2(y)$ and $f_2(x)$, respectively. Then we get the following results:

Lemma 1. *Suppose that a basis $[\mathbf{b}_1 \ldots \mathbf{b}_n]$ is chosen arbitrarily and a BKZ algorithm is performed on the basis. Suppose the same cases as in Theorem 1. Then the expectation of the first lattice vector's norm after BKZ reduction is*

$$a) \; E(\|\mathbf{b}_1\|) \lesssim 1.0447^{n-1} \cdot \det(L)^{1/n} \,. \qquad (9)$$

$$b) \; E(\|\mathbf{b}_1\|) \lesssim 1.0421^{n-1} \cdot \det(L)^{1/n} \,. \qquad (10)$$

$$c) \; E(\|\mathbf{b}_1\|) \approx 1.0157^{n-1} \cdot \det(L)^{1/n} \,. \qquad (11)$$

Proof. We use the same approach as in the proof of Theorem 1, as every BKZ reduced basis is also LLL reduced. The expectation values of the logarithms are again computed using software packages and Gaussian quadrature. For the expectation values we get -0.17496 (random $\mu_{i,i-1}$), -0.16496 (random X_i), and -0.062487 (both random). This results in the presented average-case values. $\square$

Results. Figure 4 shows the results of our BKZ analysis, i.e., it compares the norm bounds (9)(10)(11) to the experimental results of [GN08b], which is $\gamma_{\mathsf{BKZ}} \approx 1.01^n$. We notice that our average case using both random variables is more apart from the experimental results. In a BKZ reduced basis, we expect that the $\mu_{i,j}$ with $j \neq i-1$ play a more important role than in an LLL reduced basis. We have not considered those coefficients in our BKZ analysis. Including them to the analysis could lead to a lower and better bound, but is hard because of the complexity of the correlations in a BKZ reduced basis.
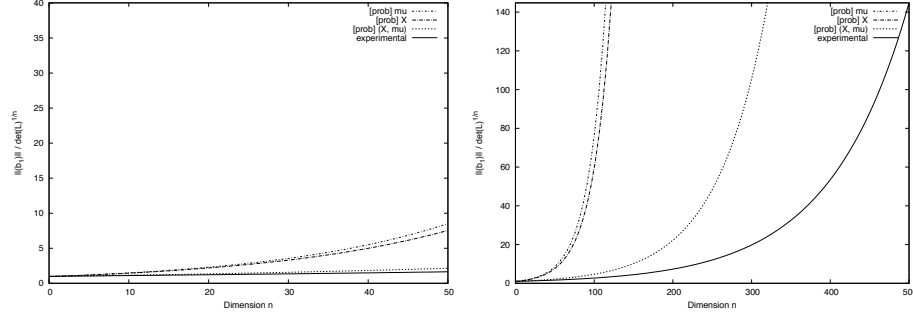

Fig. 4. Length of the first basis vector after BKZ reduction; experimental value of [NS06] and new average-case bounds.

The other results are comparable to the LLL case. The Hermite factors are lower, which is caused by the different distributions that we gained from our BKZ experiments. Again the factors are nearly the same if we use only of of both random variables.

4.3 Comparison of LLL and BKZ

Finally we compare the distribution functions of LLL and BKZ reduced bases. In the left picture of Figure 5 we have printed the distribution functions $p(y)$ and $p_2(y)$ that belong to the $\mu_{i,i-1}$ in an LLL and a BKZ reduced basis, respectively. We notice that in a BKZ reduced basis, the Gram-Schmidt coefficients $\mu_{i,i-1}$ are more concentrated in the center of the range $[-0.5, 5.5]$. This indicates that the projected basis vectors are more orthogonal than in an LLL reduced basis. That is exactly what one would guess, because BKZ reduction is a stronger notion of reducedness.

The right picture of Figure 5 shows the distribution functions $f(x)$ and $f_2(x)$, describing the distribution of the x_i's. Here the situation is a bit different. We find more very short x_i (remember that those indicate how much the Lovász is violated) in the LLL reduced basis. On the other hand, in a BKZ reduced basis the values are more spread, and values greater than 1.3 arise very seldom.
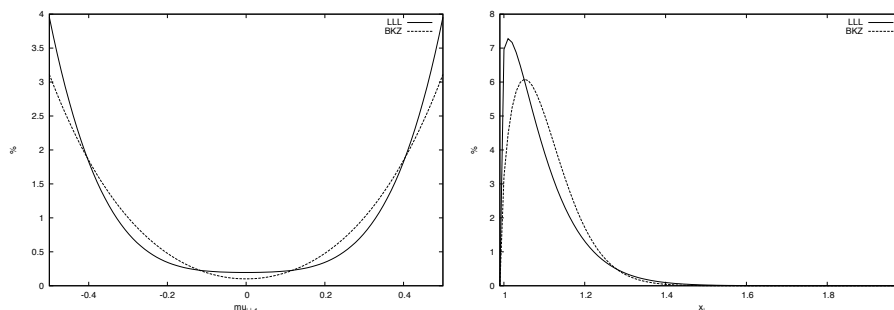
Fig. 5. Comparison of the distribution functions of $\mu_{i,i-1}$ and x_i in LLL and BKZ reduced bases.

5   Further Work

To our knowledge there is no further theoretical analysis of the deep insertion variant [SE94] concerning worst-case bounds. The best upper bound known is the standard LLL bound. It is possible to apply our probabilistic analysis to the deepLLL algorithm in order to prove some average-case bounds.

The BKZ analysis does not take all reduction properties into account. In the average-case analysis we only use the LLL properties. This has to be extended to derive better average-case estimates for BKZ. In our work we only consider reduction parameter $\delta = 0.99$, because this is the value used most in applications. For different values of δ one gets different distributions, and of course different results for the average-case bounds (as well as the worst-case bound differs for other parameters). Further on, it should be easy to adopt our Hermite factor analysis to derive average-case estimates for the approximation factor $\|\mathbf{b}_1\| / \lambda_1(L)$, but this factor is hard to control in practice and therefore was not considered in this paper.

It remains an open problem to show a more precise analysis of the Gram-Schmidt coefficients after LLL reduction and BKZ reduction, respectively. The polynomial and Weibull distributions that we assumed can only be seen as an approximation. The main difficulty in this analysis is the fact that the random variables X and Y are somehow dependent on each other. Including the dependencies into the analysis might result in a better prediction of the output norm. For this purpose it is necessary to examine the algorithm itself into detail, like [VV07] and [DFV97] do for the Gaussian algorithm and two-dimensional lattices.

Acknowledgments

We thank Markus Rückert for helpful discussions and his valuable comments.

References

[Ajt96] Miklós Ajtai. Generating hard instances of lattice problems (extended abstract). In *Proceedings of the Annual Symposium on the Theory of Computing (STOC) 1996*, pages 99–108. ACM Press, 1996.

[AKS01] M. Ajtai, R. Kumar, and D. Sivakumar. A sieve algorithm for the shortest lattice vector problem. In *Proceedings of the 33rd annual ACM Symposium on Theory of Computing*, pages 601–610. ACM Press, 2001.

[BLR08] Johannes Buchmann, Richard Lindner, and Markus Rückert. Explicit hard instances of the shortest vector problem. In *Post-Quantum Cryptography (PQCrypto) 2008*, Lecture Notes in Computer Science, pages 79–94. Springer-Verlag, 2008.

[BW02] Werner Backes and Susanne Wetzel. Heuristics on lattice basis reduction in practice. *ACM Journal of Experimental Algorithmics*, 7, 2002.

[Cas71] John W.S. Cassels. *An introduction to the geometry of numbers*. Springer-Verlag, 1971.

[CJL$^+$92] Matthijs J. Coster, Antoine Joux, Brian A. LaMacchia, Andrew M. Odlyzko, Claus-Peter Schnorr, and Jacques Stern. Improved low-density subset sum algorithms. *Computational Complexity*, 2:111–128, 1992.

[CM07] Jean-Sébastien Coron and Alexander May. Deterministic polynomial-time equivalence of computing the RSA secret key and factoring. *J. Cryptology*, 20(1):39–50, 2007.

[DFV97] Hervé Daudé, Philippe Flajolet, and Brigitte Vallée. An average-case analysis of the gaussian algorithm for lattice reduction. *Combinatorics, Probability & Computing*, 6(4):397–433, 1997.

[GGH97] Oded Goldreich, Shafi Goldwasser, and Shai Halevi. Public-key cryptosystems from lattice reduction problems. In *Advances in Cryptology — Crypto 1997*, Lecture Notes in Computer Science, pages 112–131. Springer-Verlag, 1997.

[GM03] Daniel Goldstein and Andrew Mayer. On the equidistribution of hecke points. *Forum Mathematicum 2003, 15:2*, pages 165–189, 2003.

[GN08a] Nicolas Gama and Phong Q. Nguyen. Finding short lattice vectors within mordell's inequality. In *Proceedings of the Annual Symposium on the Theory of Computing (STOC) 2008*, pages 207–216. ACM Press, 2008.

[GN08b] Nicolas Gama and Phong Q. Nguyen. Predicting lattice reduction. In *Advances in Cryptology — Eurocrypt 2008*, volume 4965 of *Lecture Notes in Computer Science*, pages 31–51. Springer-Verlag, 2008.

[Kan87] Ravindran Kannan. Algorithmic geometry of numbers. *Annual Review of Comp. Sci*, 2:231–267, 1987.

[KL78] G. A. Kabatyanski and Vladimir I. Levenshtein. *Problems of Information Transmission 14, 3*. 1978.

[Len83] Hendrik W. Lenstra. Integer programming with a fixed number of variables. *Math. Oper. Res.*, 8:538–548, 1983.

[LLL82] Arjen Lenstra, Hendrik Lenstra, and László Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 4:515–534, 1982.

[LO85]   J. C. Lagarias and Andrew M. Odlyzko. Solving low-density subset sum problems. *J. ACM*, 32(1):229–246, 1985.

[May07]   Alexander May. Using LLL-reduction for solving RSA and factorization problems, 2007. A survey for the LLL+25 conference.

[MG02]   Daniele Micciancio and Shafi Goldwasser. *Complexity of Lattice Problems: a cryptographic perspective*, volume 671 of *The Kluwer International Series in Engineering and Computer Science*. Kluwer Academic Publishers, Boston, Massachusetts, March 2002.

[NS05]   Phong Q. Nguyen and Damien Stehlé. Floating-point LLL revisited. In *Advances in Cryptology — Eurocrypt 2005*, pages 215–233, 2005.

[NS06]   Phong Q. Nguyen and Damien Stehlé. LLL on the average. In *Algorithmic Number Theory Symposium — ANTS*, pages 238–256, 2006.

[Oct]   GNU Octave. Version 3.2.2. `www.gnu.org/software/octave/`.

[Sch91]   Claus-Peter Schnorr. Factoring integers and computing discrete logarithms via diophantine approximations. In *Advances in Cryptology — Eurocrypt 1991*, pages 281–293, 1991.

[Sch94]   Claus-Peter Schnorr. Block reduced lattice bases and successive minima. *Combinatorics, Probability & Computing*, 3:507–522, 1994.

[SE94]   Claus-Peter Schnorr and M. Euchner. Lattice basis reduction: Improved practical algorithms and solving subset sum problems. *Mathematical Programming*, 66:181–199, 1994.

[Sho]   Victor Shoup. Number theory library (NTL) for C++. `http://www.shoup.net/ntl/`.

[Ste]   Damien Stehlé. Damien Stehlé's homepage at école normale supérieure de Lyon. `http://perso.ens-lyon.fr/damien.stehle/english.html`.

[VV07]   Brigitte Vallée and Antonio Vera. Lattice reduction in two dimensions: analyses under realistic probabilistic models. In *Proceedings of the 13th Conference on Analysis of Algorithms, AofA 07*. DMTCS Proceedings, 2007.