Fast Factoring Integers by SVP Algorithms, corrected

Claus Peter Schnorr
Fachbereich Informatik und Mathematik,
Goethe-Universität Frankfurt, PSF 111932,
D-60054 Frankfurt am Main, Germany.
schnorr@cs.uni-frankfurt.de
work in progress 16.06.2021

Abstract. To factor an integer \( N \) we construct \( n \) triples of \( p_n \)-smooth integers \( u, v, |u - vN| \) for the \( n \)-th prime \( p_n \). Denote such a triple a fac-relation. We get fac-relations from a nearly shortest vector of the lattice \( \mathcal{L}(\mathbf{R}_{n,f}) \) with basis matrix \( \mathbf{R}_{n,f} \in \mathbb{R}^{(n+1)\times(n+1)} \) where \( f : [1,n] \to [1,n] \) is a permutation of \([1, 2, ..., n]\) and \((f(1), ..., f(n), N' \ln N)\) is the diagonal and \((N' \ln p_1, ..., N' \ln p_n, N' \ln N)\) for \( N' = N^{\frac{3}{10}} \) is the last line of \( \mathbf{R}_{n,f} \). An independent permutation \( f' \) yields an independent fac-relation. We find sufficiently short lattice vectors by strong primal-dual reduction of \( \mathbf{R}_{n,f} \). We factor \( N \approx 2^{400} \) by \( n = 47 \) and \( N \approx 2^{800} \) by \( n = 95 \). Our accelerated strong primal-dual reduction of [GN08] factors integers \( N \approx 2^{400} \) and \( N \approx 2^{800} \) by \( 4.2 \cdot 10^9 \) and \( 8.4 \cdot 10^{10} \) arithmetic operations, much faster than the quadratic sieve and the number field sieve and using much smaller primes \( p_n \). This destroys the RSA cryptosystem.

Keywords. Primal-dual reduction, SVP, fac-relation.

1 Introduction and surview

Section 3 presents factoring algorithms for \( N \) that construct independent fac-relations from nearly shortest vectors of the lattice \( \mathcal{L}(\mathbf{R}_{n,f}) \) and quite distinct permutations \( f : [1,n] \to [1,n] \). We construct a nearly shortest vector of \( \mathcal{L}(\mathbf{R}_{n,f}) \) by primal-dual reduction of the basis \( \mathbf{R}_{n,f} \in \mathbb{R}^{(n+1)\times(n+1)} \) using \( n = 47 \) for \( N \approx 2^{400} \) and \( n = 95 \) for \( N \approx 2^{800} \) and blocks of size 24. Alg. 6.7 accelerates strong primal-dual reduction of [GN08]. This yields a nearly shortest vector of \( \mathcal{L}(\mathbf{R}_{n,f}) \). Lemma 5.1 shows that this reduction yields a fac-relation. The determinant of \( \mathbf{R}_{n,f} \) is the same for all \( f \). Independent random permutations of \([1, n]\) yield independent fac-relations. Our accelerated primal-dual reduction further halves the number of arithmetic operations. Then integers \( N \approx 2^{400} \) and \( N \approx 2^{800} \) are factored by \( 4.2 \cdot 10^9 \) and \( 8.4 \cdot 10^{10} \) arithmetic operations using a much smaller prime basis than the quadratic sieve QS and the number field sieve NFS. The main result in section 3 uses from section 5 only the upper bound (5.2) for \( M_n^T \) and (5.3). Sections 4 and 5 can be replaced by slide reduction of Gama and Nguyen [GN08] which uses no heuristics.

The enumeration algorithm Enum of [SE94] for short lattice vectors cuts stages by linear pruning. New Enum of [S94] uses the success rate \( \beta \) of stages based on the Gaussian volume heuristic. It first performs stages with high success rate and stores stages of smaller but still reasonable success rate for later performance. New Enum finds short vectors much faster than previous algorithms of Kannan [Ka87] and Fincke, Pohst [FP85] that disregard the success rate of stages. This greatly reduces the number of stages for finding a shortest lattice vector. Section 4 presents time bounds of New Enum under linear pruning for SVP for arbitrary lattice bases \( \mathbf{B} = [b_1, ..., b_n] \in \mathbb{Z}^{n \times n} \).

2 Lattices

Let \( \mathbf{B} = [b_1, ..., b_n] \in \mathbb{R}^{m \times n} \) be a basis matrix consisting of \( n \) linearly independent column vectors \( b_1, ..., b_n \in \mathbb{R}^m \). They generate the lattice \( \mathcal{L}(\mathbf{B}) = \{ \mathbf{B} \mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^n \} \) consisting of all integer linear combinations of \( b_1, ..., b_n \). The dimension of \( \mathcal{L} \) is \( n \), the determinant of \( \mathcal{L} \) is \( \det \mathcal{L} = (\det \mathbf{B} \mathbf{B}^t)^{1/2} \) for any basis matrix \( \mathbf{B} \) and its transpose \( \mathbf{B}^t \). The length of \( \mathbf{b} \in \mathbb{R}^m \) is \( ||\mathbf{b}|| = (\mathbf{b}^t \mathbf{b})^{1/2} \).

Let \( \lambda_1 = \lambda_1(\mathcal{L}) \) be the length of the shortest nonzero vector of \( \mathcal{L} \). The Hermite constant \( \gamma_n \) is the minimal \( \gamma \) such that \( \frac{2}{3} \lambda_1^2 \leq \gamma (\det \mathcal{L})^{2/n} \) holds for all lattices of dimension \( n \).

The basis matrix \( \mathbf{B} \) has the unique decomposition \( \mathbf{B} = \mathbf{Q} \mathbf{R} \in \mathbb{R}^{m \times n} \), \( \mathbf{R} = [r_{ij}]_{1 \leq i, j \leq n} \in \mathbb{R}^{n \times n} \) where \( \mathbf{Q} \in \mathbb{R}^{m \times n} \) is isometric (with pairwise orthogonal column vectors of length 1) and \( \mathbf{R} \) is
upper-triangular with positive diagonal entries $r_{i,j}$. $R = \text{GNF}(B)$ is the generic normal form of $B$. Its Gram-Schmidt coefficients $\mu_{j,i} = r_{j,i}/r_{i,i}$ are rational for integer matrices $B$. The orthogonal projection $b_i^* = \text{span}(b_1, ..., b_{i-1})^\perp$ has length $r_{i,i} = \|b_i^*\|$, $r_{1,1} = \|b_1\|$.

LLL-bases. A basis $B = QR$ is LLL-reduced or an LLL-basis for $\delta \in \{\frac{1}{2}, 1\}$ if

1. $|r_{i,j}|/r_{i,i} \leq \frac{1}{2}$ for all $j > i$ (size-reduced),
2. $\delta r_{i,i}^2 \leq r_{i+1,i+1}^2 + r_{i+1,i}^2$ for $i = 1, ..., n - 1$.

Obviously, LLL-bases satisfy $r_{i,i}^2 \leq \alpha r_{i+1,i+1}^2$ for $\alpha := 1/(\delta - \frac{1}{4})$. [LLL82] introduced LLL-bases focusing on $\delta = 3/4$ and $\alpha = 2$. A famous result of [LLL82] shows that LLL-bases for $\delta < 1$ can be computed in polynomial time and that they nicely approximate the successive minima:

3. $\alpha^{-i+1} \leq \|b_i\|^2 \lambda_i^{-2} \leq \alpha^{n-i}$ for $i = 1, ..., n$.
4. $\|b_1\|^2 \leq \alpha^n/(\det L)^{2/n}$.

A basis $B = QR \in \mathbb{R}^{m \times n}$ is an HKZ-basis (Hermite, Korkine, Zolotareff) if $|r_{i,j}|/r_{i,i} \leq \frac{1}{2}$ for all $j > i$, and if each diagonal entry $r_{i,i}$ of $B = [r_{i,j}] \in \mathbb{R}^{n \times n}$ is minimal under all transforms of $B$ to $BT$, $T \in \text{GL}_n(\mathbb{Z})$ that preserve $b_1, ..., b_{i-1}$.

A basis $B = QR \in \mathbb{R}^{m \times n}$, $R = [r_{i,j}]_{1 \leq i,j \leq n}$ is a BKZ-basis for block size $k$, (or is a BKZ-reduced) if the matrices $[r_{i,j}]_{k \leq i,j \leq k+k} \in \mathbb{R}^{2k \times 2k}$ form HKZ-bases for $k = 1, ..., n - k + 1$, see [SE94].

The efficiency of some algorithms depends on the lattice invariant $\rho(L) := \lambda_1 \gamma_n^{-1/2} (\det L)^{-1/n}$, thus $\lambda_i^2 = \rho(L)\gamma_n (\det L)^{-\frac{i-1}{n}}$. We call $\rho(L)$ the relative density of $L$. Clearly $0 < \rho(L) \leq 1$ holds for all $L$, and $\rho(L) = 1$ if and only if $L$ has maximal density. Lattices of dim $n$ of maximal density and $\gamma_n$ are known for $n = 1, ..., 8$ and $n = 24$.

3 Fast factoring integers by short vectors of the lattices $L(R_{n,f})$

Let $N > 2$ be an odd integer that is not a prime power and with all prime factors larger than $p_n$ the $n$-th smallest prime. An integer is $p_n$-smooth if it has no prime factor larger than $p_n$. The classical method factors $N$ by $n + 1$ independent pairs of $p_n$-smooth integers $u, |u - v|N$. We call such $u, |u - v|N$ a fac-relation. Our factoring method generates fac-relations with $p_n$-smooth $v$.

The classical method of factoring $N$. Given $n + 1$ fac-relations $(u_j, v_j)$ we have for $p_0 := -1$

$$u_j = \prod_{i=0}^{n} p_i^{e_{i,j}}, \quad u_j - v_j N = \prod_{i=0}^{n} p_i^{e_{i,j}}$$

with $e_{i,j}, e'_{i,j} \in \mathbb{N}$. (3.1)

We have $(u_j - v_j N)/u_j \equiv 1 \mod N$ since $(u_j - v_j N) = u_j \mod N$. Hence

$$\prod_{i=0}^{n} p_i^{e_{i,j} - e'_{i,j}} \equiv 1 \mod N.$$  

Any solution $t_1, ..., t_{n+1} \in \{0, 1\}$ of the equations

$$\sum_{j=1}^{n+1} t_j (e_{i,j} - e'_{i,j}) \equiv 0 \mod 2$$

solves $X^2 - 1 = (X - 1)(X + 1) \equiv 0 \mod N$ by $X = \prod_{i=0}^{n} p_i^{\frac{1}{2} \sum_{j=1}^{n+1} t_j (e_{i,j} - e'_{i,j})} \mod N$. If $X + 1 \equiv 0 \mod N$ this yields two non-trivial factors $\gcd(X \pm 1, N) \neq \{1, N\}$ of $N$.

The linear equations (3.2) can be solved within $O(n^3)$ bit operations. We neglect this minor part of the work load of factoring $N$. Hence $N$ can be factored by finding about $n + 1$ fac-relations. This factoring method goes back to Morrison & Brillhart [MB75] and let to the first factoring algorithm in subexponential time by J. Dixon [DS81].

We generate fac-relations from short vectors of the lattices $L(R_{n,f})$ where $f : [1, n] \rightarrow [1, n]$ is a permutation of $[1, n] = [1, 2, ..., n]$. We construct short vectors of $R_{n,f} \in \mathbb{R}^{(n+1) \times (n+1)}$ by strong primal-dual reduction with algorithm 3.2. In order to get distinct fac-relations from distinct permutations $f : [1, n] \rightarrow [1, n]$ it is important that these permutations are quite different, for instance nearly random. The first $n$ lines of $R_{n,f}$ have all nonzero entries on the diagonal.

$$R_{n,f} = \begin{bmatrix} f(1) & 0 & 0 & \cdots & 0 & f(n) \\ 0 & \ddots & \vdots & \cdots & \vdots & 0 \\ \cdots & \vdots & 0 & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & 0 \\ N' \ln p_1 & \cdots & N' \ln p_n & N' \ln N \end{bmatrix} = [b_1, ..., b_{n+1}]$$
Here \( \ln = \log_e \) for the Euler number \( e = 2.7182818284 \cdots \). Let \( N' = N^{1/(n+1)} \) and \( \mathbf{R}_{n,f}' = [\mathbf{b}_1, \ldots, \mathbf{b}_l] \). We identify each vector \( \mathbf{b} = \sum_{i=1}^l e_i \mathbf{b}_i \in \mathcal{L}(\mathbf{R}_{n,f}') \) with the pair \((u,v)\) of relative prime and \( p_n\)-smooth integers

\[
u = \prod_{e_i>0} p_i^{e_i}, \quad v = \prod_{e_i<0} p_i^{-e_i} \in \mathbb{N} \quad \text{denoting} \quad \mathbf{b} \sim (u,v).
\]

For \( \mathbf{b} \sim (u,v) \) we denote \( \hat{z}_n := N' \ln \frac{u}{v}, \hat{z}_{b-b_{n+1}} := N' \ln \frac{2}{\ln v} \) the last coordinates of \( \mathbf{b} \) and \( \mathbf{b} - \mathbf{b}_{n+1} \). As a factor \( p_i^{e_i} \) of \( uv \) adds \( e_i \ln p_i \) to \( \ln uv \) and \( e_i^2 \ln p_i \) to \( \|b\|^2 \) we have \( \|b\|^2 \geq \ln uv + \hat{z}_n^2 \) with equality if and only if \( uv \) is squarefree so that \( e_i \in \{ -1, 0, 1 \} \) for all \( i \). Similarly

\[
\|b - b_{n+1}\|^2 \geq \ln uv + \hat{z}_{b-b_{n+1}}^2 \quad \text{holds for} \quad (u,v) \in \mathcal{L}(\mathbf{R}_{n,f}').
\]

**Lemma 3.1** We have \( \hat{z}_{b-b_{n+1}} = N' \ln(\frac{2}{\ln v}) = -N' \sum_{i=1}^l (-x)^j/i \) for \( x = \frac{2}{\ln v} \), \((u,v) \sim \mathbf{b} \in \mathcal{L}(\mathbf{R}_{n,f}'). \) Let \( x \in [-\frac{1}{2}, \frac{1}{2}] \) and \( \|b - b_{n+1}\| = \lambda_3(\mathcal{L}(\mathbf{R}_{n,f}')) \) then \( |u - vN| < v|\hat{z}_{b-b_{n+1}}|/(1-\varepsilon/2) \) holds for \( 0 < \varepsilon < \frac{1}{2} \) if either \( vN < u < (1+\varepsilon)vN \) or \( u < vN < (1+\varepsilon)u \).

**Proof.** We apply the Taylor form \( \ln(1 + x) = \sum_{j=1}^\infty (-x)^j/j \) holding for \( x \in [-\frac{1}{2}, 1] \). Clearly \( \hat{z}_{b-b_{n+1}} \) lies between the sums \( -N' \sum_{i=1}^l (-x)^j/i \) for \( j = 1, 2 \).

If \( vN < u < (1+\varepsilon)vN \) then \( \frac{N'\ln(\frac{2}{\ln v})}{vN} < \hat{z}_{b-b_{n+1}} \) and this implies \( u - vN' < v|\hat{z}_{b-b_{n+1}}|/(1-\varepsilon/2) \).

If \( u < vN < (1+\varepsilon)u \) then \( \frac{N'\ln(\frac{2}{\ln v})}{vN} < \hat{z}_{b-b_{n+1}} \) and this implies \( v - u < v|\hat{z}_{b-b_{n+1}}|/(1-\varepsilon/2) \).

**Lemma 3.1** shows \( |u - vN'| = p_n^k \) for \( z \in \ln |u - vN'|/\ln p_n \). Hence random \( |u - vN| \) is \( p_n\)-smooth and yields a fac-relations with probability \( \rho(z) \), \( \rho(z) \) is the Dickman, de Bruin-p-function, see [G06].

If \( z = \lfloor z \rfloor + \hat{z} \), with \( 0 < \hat{z} < 1 \) then \( \rho(z) \approx \rho\lfloor z \rfloor \left( \frac{\hat{z}}{\rho\lfloor z \rfloor} \right) \). Note that large \( f(i) \) implies that \( p_i \) is unlikely a factor of \((u,v)\) of the constructed fac-relations \( u, v, |u - vN'| \). For quite different permutations \( f, f' \) this implies that they generate different fac-relations.

**Algorithm 3.2** for lattice reduction of \( \mathcal{L}(\mathbf{R}_{n,f}') \)

1. **LLL-reduce** \( \mathbf{R}_{n,f} \) for \( \alpha = 1/(\delta - \frac{1}{2}) \), compute \( \mathbf{R} = \text{GNF}(\mathbf{R}_{n,f}) \in \mathbb{R}^{(n+1)\times (n+1)} \) in pol. time.

2. **Primal-dual reduce** \( \mathbf{R} \) to \( \mathbf{R}^\prime \) following [GHK06] by algorithm 6.3 and iteratively increasing the block size following [AWHTT16]. This yields a vector \( \mathbf{b}_1 \in \mathcal{L}(\mathbf{R}) \) satisfying

\[
\|b_1\|^2 \leq \gamma_k(\alpha^2)^{\frac{n}{n+1}} (\det \mathbf{R})^{\frac{1}{n+1}}.
\]

**Number of arithmetic operations**

**Algorithm 3.2** for \( N \approx 2^{400} \) and factoring \( N \approx 2^{400} \):

For \( n + 1 = 48 \) we have \( \text{det}(\mathbf{R}_{n,f}')) \frac{3}{2} = 47! N' \ln N \approx 480.67 \). Primal-dual reduce the basis \( \mathbf{R}_{47,f} \in \mathbb{R}^{48 \times 48} \) by alg. 3.2 where \( 48 = hk, k = 24, h = 48/k = 2 \). Theorem 6.4 shows that step 2 of algorithm 3.2 performs at most \( 48^k \cdot \log_2(\alpha) \) iterations. Each iteration HKZ-reduces two blocks \( p_{b_{k+1}}, \mathbf{R}' \in \mathbb{R}^{k \times k} \) performing per block \( k^{k/4+1} \) arithmetic operations according to (5.3), and activates this reduction if a subsequent size-reduction of the columns \( \{b_{k+1}, \ldots, b_{k+k}\} \) of the GNF decreases \( \text{det}(\mathbf{R}_k) \) by the factor \( \delta_{b_k}' \). Step 2 of alg. 3.2 performs at most \( 48^k \cdot \log_2(\alpha) k^{k/4+1} \) arithmetic operations, where \( \log_2(\alpha) = 1 \). The minor work for LLL-reduction can be neglected. Alg. 3.2 is performed 48 times to find 48 fac-relations. This requires at most \( 48 \cdot 1.75 \cdot 10^8 = 8.4 \cdot 10^9 \) arithmetic operations.

**Algorithm 3.2** for \( N \approx 2^{400} \) and factoring \( N \approx 2^{400} \):

For \( n + 1 = 96 \) we have \( \text{det}(\mathbf{R}_{n,f}')) \frac{3}{2} = 95! N' \ln N \approx 2289.44 \). Primal-dual reduce the basis \( \mathbf{R}_{55,f} \in \mathbb{R}^{96 \times 96} \) by algorithm 6.3 where \( 96 = hk, k = 24, h = 96/k = 4 \). By theorem 6.2 this yields a vector \( \mathbf{b}_1 \in \mathcal{L}(\mathbf{R}_{55,f}) \) with \( \|b_1\|^2 \leq \gamma_k(\alpha^2)^{\frac{n}{n+1}} (\det(\mathbf{R}_{55,f}')) \frac{3}{2} < 0.8408696 \). Hence \( \|b_1\| < 0.917 \). Lemma (3.1) shows for \( b = b_{n+1} \sim (u,v), v \leq p_n = p_{b_{55}} = 499, \varepsilon = \frac{1}{2} \) that

\[
|u - vN| < p_n |\hat{z}_{b-b_{n+1}}|/(1-\varepsilon/2) \leq 522.946 < p_n^{0.007544}
\]
\[ \rho(0.007544) = \rho(2)^{0.007544} = 0.991126 \] and \( |u - v| \) is \( p_n \)-smooth and yields fac-relations with prob. \( \approx 1 \). Then alg. 6.3 and alg. 3.2 each perform \( \frac{9624k^{k/8+1.1}}{\alpha} \approx 1.4 \cdot 10^9 \) arithmetic operations for \( \alpha = 1/\delta_b \). To find 96 fac-relations by iterating alg. 3.2 this performs at most 96 \( 1.4 \cdot 10^9 = 1.344 \cdot 10^{11} \) arithmetic operations.

**Using strong primal-dual reduction of Gama, Nguyen [GN08]** based on the accelerated alg. 6.7 performs about half as many arithmetic operations as alg. 6.3 for primal-dual reduction. It factors integers \( N \approx 2^{400} \) and \( N \approx 2^{800} \) by \( 4.2 \cdot 10^9 \) and \( 8.4 \cdot 10^{10} \) arithmetic operations.

Hence the number of arithmetic operations for factoring \( N \) increases from \( N \approx 2^{400} \) to \( N \approx 2^{800} \) by the factor 20. Again it increases from \( N \approx 2^{800} \) to \( N \approx 2^{1200} \) and to \( N \approx 2^{1600} \) by the factor 20. Hence \( N \approx 2^{400} \) can be factored by \( 4.2 \cdot 10^9 \) arithmetic operations in about 1 minute and \( N \approx 2^{1200} \) can be factored by about \( 20^3 = 800 \) minutes or in about 5.55 days.

**Using Improved Progressive BKZ Algorithm of [AWHTT16]** can still accelerate the time for finding shortest lattice vectors, in particular for factoring \( N \approx 2^{800} \) and finding a nearly shortest vector in a lattice with basis \( \mathbb{R}_{5,f} \in \mathbb{R}^{96 \times 96} \). It may also be helpful to use the results of [MW16] to speed up lattice reduction.

**Factoring time bounds for quadratic sieve QS and number field sieve NFS**: The QS uses for the factoring of \( N \approx 2^{400} \) that \( p_n \approx e^{1/2\sqrt{\ln \ln N}} \approx 3.76 \cdot 10^9 \), see [CP01, section 6.1]. The prime base for NFS is bigger than for QS. The number of arithmetic steps of our factorisation is quite small compared with QS and NFS factorisation but the bit length of integers is large. The numbers of arithmetic operations for QS, NFS factorisation of \( N \approx 2^{400} \) in [CP01, section 6.2] are:

\[
\begin{align*}
\rho^{\sqrt{\ln N \ln \ln N}} & \approx 1.415 \cdot 10^{17} & \text{for QS} \\
\rho^{(64/9)^{1/3}(\ln N)^{1/3}(\ln \ln N)^{2/3}} & \approx 1.675 \cdot 10^{17} & \text{for NFS}.
\end{align*}
\]

**NFS factoring of \( N \approx 2^{800} \) performs** \( 2.8126 \times 10^{23} \) arithmetic operations.

### 4 Efficient enumeration of short lattice vectors

We outline the SVP-algorithm based on the success rate of stages. **New Enum** improves the algorithm Enum of [SE94, SH95]. We recall Enum and present New Enum as a modification that essentially performs all stages of Enum in decreasing order of success rates. This SVP-algorithm New Enum finds a shortest lattice vector fast without enumerating all shortest lattice vectors.

Let \( B = [b_1, \ldots, b_n] = QR \in \mathbb{R}^{m \times n}, R = [r_{i,j}]_{1 \leq i, j \leq n} \in \mathbb{R}^{n \times n} \), be the given basis of \( L = L(B) \). Let \( \pi_t \) : \( \text{span}(b_1, \ldots, b_n) \to \text{span}(b_1, \ldots, b_{t-1}) \) be the orthogonal projections and let \( L_t = L(b_1, \ldots, b_{t-1}) \).

**The success rate of stages.** At stage \( u = (u_1, \ldots, u_n) \) of Enum for SVP of \( L \) a vector \( b = \sum_{i=1}^{\infty} u_i b_i \in L \) is given such that \( \| \pi_t(b) \|^2 \leq \lambda_t^2 \). (When \( \lambda_t^2 \) is unknown we use instead some \( \lambda > \lambda_t^2 \)) Stage \( u \) calls the substages \( (u_{t-1}, \ldots, u_n) \) such that \( \| \pi_{t-1}(\sum_{i=t-1}^{\infty} u_i b_i) \|^2 \leq \lambda_t^2 \). We have \( \| \sum_{i=1}^{\infty} u_i b_i \|^2 = \| \zeta + \sum_{i=1}^{t-1} u_i b_i \|^2 + \| \pi_t(b) \|^2 \), where \( \zeta := b - \pi_t(b) \in \text{span}(L_t) \) is \( b \)'s orthogonal projection in span \( L \). Stage \( u \) and its substages enumerate the intersection \( B_{t-1}(\zeta, \varrho_l) \cap L_t \) of the sphere \( B_{t-1}(\zeta, \varrho_l) \subset \text{span} L_t \) with radius \( \varrho_l := (\lambda_t^2 - \| \pi_t(b) \|^2)^{1/2} \) and center \( \zeta \). The Gaussian volume heuristics for \( t = 1, \ldots, n \) the expected size \( |B_{t-1}(0, \varrho_l) \cap (\zeta + L_t)| \) to be the success rate

\[
\beta_l(u) = \Delta_t \text{vol} B_{t-1}(0, \varrho_l) / \text{det} L_t \tag{4.1}
\]

standing for the probability that there is an extension \( (u_1, \ldots, u_n) \) of \( u = (u_1, \ldots, u_n) \) such that \( \| \sum_{i=1}^{\infty} u_i b_i \| \leq \lambda_t \). Here vol \( B_{t-1}(0, \varrho_l) = V_{t-1} \varrho_l^{t-1} \). \( V_{t-1} = \frac{\pi^{t-1} \varrho_l^{t-1}}{(t-1)!} \approx \left( \frac{\pi}{2t-2} \right)^{t-1} / \sqrt{\pi(t-1)} \) is the volume of the unit sphere of dimension \( t-1 \) and det \( L_t = \text{det} L_{t-1} \cdots \text{det} L_1 \). If \( \zeta \in \text{span} L_t \) is uniformly distributed the expected size of this intersection satisfies \( E_{\zeta} \left[ \#(B_{t-1}(0, \varrho_l) \cap (\zeta + L)) \right] = \frac{\pi^{t-1} \varrho_l^{t-1}}{(t-1)!} \sqrt{\pi(t-1)} \)
\( \beta_t(u) \). This holds because \( 1/\det L_t \) is the number of lattice points of \( L_t \) per volume in span \( L_t \). We do not simply cut \( u_t \) due to a small \( \beta_t \) because there might be a vector in \( L_t \) very close to \( \zeta_t \).

The success rate \( \beta_t \) has been used in [SH95] to speed up ENUM by cutting stages of very small success rate. New Enum first performs all stages with sufficiently large \( \beta_t \) giving priority to small \( t \) and collects during this process the unperformed stages in the list \( L \). For instance it first performs all stages with \( \beta_t \geq 2^{-\varphi} \log_2(t) \). Thereafter New Enum increases \( s \) to \( s + 1 \). So far our experiments simply perform all stages with \( \beta_t \geq 2^{-\varphi} \). If \( \lambda_1^t \) is unknown we can compute \( g_t, \beta_t \) replacing \( \lambda_1^t \) by the upper bound \( A = \frac{1.744}{n \log n} \det(B(B)) \geq \lambda_1^t \) which holds since \( \gamma_n \leq 1.744 \) in \( n \) holds for \( n \geq n_0 \) by a computer proof of Kabatiansky, Levenshtein [KaLe78].

Dichotomy for SHORT VECTORS

For short vectors \( b \), \( b \), \( \beta_t \) has been used in [SH95] to speed up ENUM and cuts the unperformed stages from the list \( L \).

\[ 2.1. \]

Let \( L \) be a sequence of stages. For the current \( b \) of block size 32, \( A \), \( s = \log n = \log_2 n \).

\[ 2.2. \]

\[ \text{NEW ENUM} \]

\[ \text{INPUT} \] BKZ-basis \( B = QR, R = [r_{i,j}] \in \mathbb{R}^{n \times n} \) of block size 32, \( A, s = \log n = \log_2 n \).

\[ \text{OUTPUT} \] a sequence of \( b \in \mathcal{L}(B) \) of decreasing length terminating with \( \|b\| = \lambda_1 \).

1. \( L = \emptyset \).

2. \( L \) is not empty.

\[ \text{NEW ENUM} \]

\[ \text{OUTPUT} \] a sequence of \( b \in \mathcal{L}(B) \) such that \( \|b\| \) decreases to \( \lambda_1 \).

1. \( L := \emptyset \), \( t := t_{\text{max}} := 1, \) \( \forall \) \( i = 1, \ldots, n \). \( D \) \( c_i := u_i := y_i := 0, \) \( \nu_i := u_i := 1, \) \( s := 5, \) \( c_1 := r_{1,1}^2 \).

2. \( s := s + 1, \) \( \text{IF} \) \( L \neq \emptyset \) \( \text{THEN} \) perform all stages \( u_i \in L \) with \( \beta_t(u_i) \geq 2^{-s} \log_2 t \).

Running in linear space. If instead of storing the list \( L \) we restart NEW ENUM in step 3 on level \( s + 1 \) then NEW ENUM runs in linear space and its running time increases at most by a factor \( n \).

Practical optimization. NEW ENUM computes \( R, \beta_t, V_t, g_t, c_t \) in floating point and \( b \), \( \|b\|^2 \) in exact arithmetic. The final output \( b \) has length \( \|b\| = \lambda_1 \), but this is only known when the more expensive final search does not find a vector shorter than the final \( b \).

Reason of efficiency. For short vectors \( b = \sum_{i=1}^n u_i b_i \in \mathcal{L}(0) \) the stages \( u = (u_1, \ldots, u_n) \) have large success rate \( \beta_t(u) \). On average \( \|\pi_t(b)\|^2 \approx \frac{1}{n^2} \lambda_1^2 \) holds for a random \( b \in \mathbb{R}^n \) of length \( \lambda_1 \). Therefore \( g_t^2 = A - ||\pi_t(b)||^2 \) and \( \beta_t(u) \) are large. NEW ENUM tends to output very short lattice vectors first.

New Enum for SVP

\[ \text{INPUT} \] BKZ-basis \( B = QR, R = [r_{i,j}] \in \mathbb{R}^{n \times n}, A \geq \lambda_1^2, s_{\text{max}} \)

\[ \text{OUTPUT} \] a sequence of \( b \in \mathcal{L}(B) \) such that \( \|b\| \) decreases to \( \lambda_1 \).

1. \( L := \emptyset, t := t_{\text{max}} := 1, \) \( \forall \) \( i = 1, \ldots, n \). \( D \) \( c_i := u_i := y_i := 0, \) \( \nu_i := u_i := 1, \) \( s := 5 \).

2. \( s := s + 1, \) \( \text{IF} \) \( L \neq \emptyset \) \( \text{THEN} \) perform all stages \( u_i \in L \) with \( \beta_t(u_i) \geq 2^{-s} \log_2 t \).

3. \( s := s + 1, \) \( \text{IF} \) \( s > s_{\text{max}} \) \( \text{THEN} \) restart with a larger \( s_{\text{max}} \).
New Enum is particularly fast for small $\lambda_i$. The size of its search space approximates $\lambda_i^2 V_n$, and is by Prop. 4.1 heuristically polynomial if $rd(L) = o(n^{-1/4})$. Having found $b'$ New Enum proves $\|b'\| = \lambda_i$ in exponential time by a complete exhaustive enumeration.

**Notation.** We use the following function $c_t: \mathbb{Z}^{n-t+1} \to \mathbb{R}$:
\[
c_t(u_1, \ldots, u_n) = \|\pi_t((\sum_{i=1}^n u_i b_i))\|^2 = \sum_{i=1}^n (\sum_{j=1}^t u_i r_{i,j})^2.
\]

Hence
\[
c_t(u_1, \ldots, u_n) = (\sum_{i=1}^n u_i r_{i,t+1})^2 + c_{t+1}(u_{t+1}, \ldots, u_n).
\]

Given $u_t+1, \ldots, u_n$ Enum takes for $u_t$ the integers that minimize $|u_t + y_t|$ for $y_t := \sum_{i=t+1}^n u_i r_{i,t}/r_{i,t}$ in order of increasing distance to $-y_t$ adding to the initial $u_t := -|y_t|$ iteratively $|\nu_t/2|(-1)^\nu t$.

where $c_t := \text{sign}(u_t + y_t) \in \{\pm 1\}$ and $\nu_t$ numbers the iterations starting with $\nu_t = 0, 1, 2, \ldots$:
\[
-\lfloor y_t \rfloor, -\lfloor y_t \rfloor - 1, -\lfloor y_t \rfloor + 1, -\lfloor y_t \rfloor - 2, -\lfloor y_t \rfloor + 2, \ldots, -\lfloor y_t \rfloor + |\nu_t/2|(-1)^{\nu t}, \ldots,
\]

where $\text{sign}(0) := 1$ and $[r]$ denotes a nearest integer to $r \in \mathbb{R}$. The iteration does not decrease $|u_t + y_t|$ and $c_t(u_t, \ldots, u_n)$, it does not increase $\nu_t$ and $\beta_t$. Enum performs the stages $(u_t, \ldots, u_n)$ for fixed $u_{t+1}, \ldots, u_n$ in order of increasing $c_t(u_t, \ldots, u_n)$ and decreasing success rate $\beta_t$. $\beta_t$ extends this priority to stages of distinct $t, t'$ taking into account the size of two spheres of distinct dimensions $n - t, n - t'$.

The center $c_t = b - \pi_t(b) = \sum_{i=t}^n u_i (b_i - \pi_i(b_i))$ in $\mathcal{L}(L_t)$ changes continuously within New Enum which improves Enum when step 3 performs stages $u_t \ast$ in $L$ the current $u_t$ and this can make the stored $\beta_t$ of $u_t$ smaller than $2^{-\delta}$ so that $u_t \ast$ will not be performed but must be stored in $L$ with the adjusted smaller values $\beta_t$, $\beta_t \ast$. The stored stages $u_t \ast$ with $\beta_t \ast \geq 2^{-\delta}$ should be performed in a succession giving priority to large success rates and small $t$.

**Time for solving SVP for $L(B)$, New Enum.** New Enum performs for each $s = 5, 6, \ldots, s_{\text{max}}$ only stages $u_t$ with success rate $\beta_t \geq 2^{-s}$. Let $#b_{s,A}$ denote the number of performed stages with $t, s, A$. If $\beta_t$ is a reliable probability then New Enum performs on average at most $2^s$ stages with success rate $\beta_t \geq 2^{-s}$ before decreasing $A$ - this number of performed stages is even smaller than $2^s$ since New Enum also performs stages with success rate $\beta_t \geq 2^{-s-1}$. New Enum performs for each stage of step 2 on average at most $2(n-t)(1-o(1))$ arithmetical steps for computing $y_t$ which add up to $\sum_{i=1}^n 2(n-t)(1+o(1)) \approx n(n+1)(1+o(1))$ arithmetical steps and it performs $O(n)$ arithmetical steps for testing that $\ln \beta_t \geq -n \ln 2$ for $t = 1, \ldots, n$ using $\beta_t \approx V_{t-1} \rho_{t-1}^{-1}/d \mathcal{L}$ assuming that $\ln(2\pi n) \ln n(1+x)$ for $x = 1, \ldots, n$ are given for free.

If the initial basis $B \in \mathbb{R}^{n \times n}$ is a BKZ-basis with block size $k$ then $\|b_{11}\| \leq \lambda_1^{\frac{n-1}{k}}$. As New Enum performs stages with high success rates first then each decrease of $A$ will on average halve $A/\lambda_i^2$ so that there are at most $\log_2(A/\lambda_i^2)$ iterations of step 2 that decrease the initial $A$ of step 1. So after the initial reduction of $B$ New Enum solves SVP for $s_{\text{max}}$ with error probability $o(1)$ and performs on average at most $O(n^{2s_{\text{max}}})$ arithmetical steps for each $A$. Hence SVP is solved by
\[
2^{s_{\text{max}}}(n^2 + O(n))2^{\frac{n-1}{k} \log_2 \gamma_k} \text{ arithmetical steps. (4.2)}
\]

## 5 New Enum for SVP with linear pruning

The heuristics of linear pruning gives weaker results but is easier to justify than handling the success rate $\beta_t$ as a probability function. Proposition 5.1 bounds under linear pruning the time to find $b' \in \mathcal{L}(B)$ with $\|b'\| = \lambda_1$. It shows that SVP is polynomial time if $rd(L)$ is sufficiently small. Note that finding an unproved shortest vector $b'$ is easier than proving $\|b'\| = \lambda_1$. New Enum finds an unproved shortest lattice vector $b'$ in polynomial time under the following conditions and assumptions:

- the given lattice basis $B = [b_1, \ldots, b_n]$ and the relative density $rd(L)$ of $\mathcal{L}(B)$ satisfy $rd(L) \leq (\frac{\sqrt{n-1}}{\pi n} \lambda_1^{\frac{n-1}{n}})^2$, i.e., both $b_1$ and $rd(L)$ are sufficiently small.

**GSA:** The basis $B = QR, R = [r_{i,j}]_{1 \leq i, j \leq n}$ satisfies $r_{i,j}^2/r_{i-1,j-1}^2 = q$ for $2 \leq i \leq n$ for some $q > 0$.

**SA:** There is a vector $b' \in \mathcal{L}(B)$ such that $\|b'\| = \lambda_1$ and $\|\pi_t(b')\|^2 \leq \frac{n-t+1}{n} \lambda_t^2$ for $t = 1, \ldots, n$.

(Later we will use a similar assumption CA for CVP).

- the vol. circ. is close: $M''_t := \#B_{n-t+1}(0, \rho_t) \cap \pi_t(L) \approx \frac{\sum_{k=0}^{n-t+1} \rho_t^{n-t+1} \Gamma(\pi_t(L))}{\delta t \pi_t(L)}$ for $\rho_t = \frac{n-t+1}{n} \lambda_t^2$.
Remarks. 1. If GSA holds with $q \geq 1$ the basis $B$ satisfies $\|b_i\| \leq \frac{1}{\sqrt{n}} + \frac{3}{\sqrt{n}} \lambda_i$ for all $i$ and $\|b_i\| = \lambda_i$. Therefore, $q < 1$ unless $\|b_i\| = \lambda_i$. GSA means that the reduction of the basis is "locally uniform", i.e., the $r_i^2$ form a geometric series. It is easier to work with the idealized property that all $r_{i,j} = r_{i,j-1} = \lambda_j$. In practice $r_{i,j} = r_{i,j-1} = \lambda_j$ slightly increases on the average with $i$. [BL05] studies "nearly equality". GSA has been used in [S03, S06, GN06, S010, N10] and in the security analysis of NTRU in [H07, HHH2009].

2. The assumption SA is supported by a fact proven in the full paper of [GNR10]:

$$\Pr[\|\pi_i(b')\|^2 \leq \frac{n-t+1}{n} \lambda_i^2 \text{ for } t = 1, \ldots, n] = \frac{1}{n}$$

for random $b' \in \text{Span}(L)$ with $\|b'\| = \lambda_1$.

Linear pruning means to cut off all stages $(u_t, \ldots, u_n)$ that satisfy $\|\pi_i(\sum_{i=1}^n u_i b_i)\|^2 > \frac{n-t+1}{n} \lambda_i^2$. Linear pruning is impractical because it does not provide any information on SVP in case of failure. We use linear pruning only as a theoretical model for easy analysis. We have implemented SVP via New Enum and we will show in section 5 that stages $(u_t, \ldots, u_n)$ that are cut by linear pruning have extremely low success probability so they will not be performed by New Enum.

3. Errors of the volume heuristics. The minimal and maximal values of $\#_n := \#(B(\zeta_n, \rho_n) \cap L)$, and similar for $\#_t := \#(B(\zeta_n, \rho_n) \cap L_t)$, are for fixed $n, \rho_n$ very close for large radius $\rho_n$, but can differ considerably for small $\rho_n$ since $\#_n$ can change a lot with the actual center $\zeta_n$ of the sphere. For small $\rho_n$ the minimum of $\#_n$ can be very small and then the average value for random center $\zeta_n$ is closer to the maximum of $\#_n$. For more details see the theorems and Table 1 of [MO90]. As New Enum works with average values for $\#_n$, $\#_t$ its success rate $\beta_t$ frequently overestimates the success rate for the actual $\zeta_n$. A cut of the smallest (resp. closest) lattice vector by New Enum in case that it underestimates $\#_t$ can nearly be excluded if stages are only cut for very small $\beta_t$.

Our time bounds must be multiplied by the work load per stage, a modest polynomial factor covering the steps performed at stage $(u_t, \ldots, u_n)$ of Enum before going to a subsequent stage.

Proposition 5.1 Let the basis $B = QR$, $R \in \mathbb{R}^{n \times n}$ of $L$ satisfy $rd(L) \leq \left(\frac{1}{\sqrt{n}} + \frac{\sqrt{\pi}}{\sqrt{2n}}\right)^{\frac{1}{2}}$ and GSA and let $L$ have a shortest lattice vector $b'$ that satisfies SA. Then Enum with linear pruning finds such $b'$ under the volume heuristic in polynomial time.

Proof. For simplicity we assume that $\lambda_1$ is known. Pruning all stages $(u_t, \ldots, u_n)$ that satisfy $\|\pi_i(\sum_{i=1}^n u_i b_i)\|^2 > \frac{n-t+1}{n} \lambda_i^2 := \rho_i^2$ does not cut off any shortest lattice vector $b'$ that satisfies SA.

The volume heuristics approximates the number $M_t^n$ of performed stages $(u_t, \ldots, u_n)$ to

$$M_t^n = \#B_{n-t+1}(0, \rho_t) \cap \pi_t(L) \approx \left(\frac{n-t+1}{n} \lambda_t\right)^{n-t+1} V_{n-t+1}/(r_{t,1} \cdots r_{t,n})$$

$$\approx \left(\lambda_t \left(\frac{2n}{\sqrt{n}}\right)^{n-t+1}/(r_{t,1} \cdots r_{t,n}) \right)^{n-t+1}/(r_{t,1} \cdots r_{t,n}) = \left(\lambda_t \left(\frac{2n}{\sqrt{n}}\right)^{n-t+1}/(r_{t,1} \cdots r_{t,n}) \right)^{n-t+1}/(r_{t,1} \cdots r_{t,n})$$

(5.1)

Here $\approx$ uses Stirling’s approximation $V_n = \pi^{n/2}/(n/2)! \approx (\frac{2n}{\sqrt{n}})^{n/2}/\sqrt{\pi n}$. Obviously $\|b_t\| = r_{1,t}q^{n-t}$ holds by GSA and thus

$$r_{t,1} \cdots r_{t,n} \leq q^{n-t} = q^{n-t}$$

For $t = 1$ this yields $q^{n-t} = (\det L)^t/(\lambda t) = 1/\lambda_t (\sqrt{\pi rd(L)})^t$. Combining (5.1) with this equation and $\gamma_n < \frac{n}{\lambda_1}$ which holds for $n > n_0$, we get

$$M_t^n \leq \left(\frac{\lambda_t \left(\frac{2n}{\sqrt{n}}\right)^{n-t+1}/(r_{t,1} \cdots r_{t,n}) \right)^{n-t+1}/(r_{t,1} \cdots r_{t,n})$$

(5.2)

Evaluating this upper bound for $rd(L) \leq \left(\frac{1}{\sqrt{n}} + \frac{\sqrt{\pi}}{\sqrt{2n}}\right)^{\frac{1}{2}}$ yields

$$M_t^n \leq \left(\frac{1}{\sqrt{n}} + \frac{\sqrt{\pi}}{\sqrt{2n}}\right)^{\frac{1}{2}} \frac{1}{\lambda_t} = \left(\frac{1}{\sqrt{n}} + \frac{\sqrt{\pi}}{\sqrt{2n}}\right)^{\frac{1}{2}} \frac{1}{\lambda_t} \left(\frac{1}{\sqrt{n}} + \frac{\sqrt{\pi}}{\sqrt{2n}}\right)^{\frac{1}{2}} \frac{1}{\lambda_t}$$

This approximate upper bound has for $t \leq n$ its maximum 1 at $t = n$. This proves Prop. 5.1. □

Note that (5.2) only assumes the volume heuristic and GSA, but no upper bound on $rd(L)$.
SVP-time bound for $rd(L) \leq 1$ under linear pruning. (5.2) proves for $rd(L) \leq 1$ that

$$M_t^f \lesssim \left( \frac{r^{1+t} \ln^2 n}{\lambda^4} \right)^{n^{-\frac{(t-1)(t+2)}{n-1}} n^{-\frac{t+1}{2n-1}}}.$$  

The exponent $n - \frac{(t-1)(t+2)}{n-1} - n - t - 1$ is maximal for $t = n/2 + 1$ with maximal value $\frac{3}{8} n^2$. This proves for $r_{1,1}/\lambda_1 = n^{o(1)} \sqrt{e/\pi}$. The heuristic SVP time bound

$$O(n) \left( \frac{r^{1+t} \ln^2 n}{\lambda^4} \right)^{\frac{3}{8} n^2/2^t} = n^{n/8+1.1}.$$  

This beats under heuristics the proven SVP time bound $n^{n/2+o(n)}$ of Hanrot, Stehle [HS07] which holds for a quasi-HKZ-basis $B$ satisfying $\|b_1\| \leq 2\|b_2\|$ and having a HKZ-basis $\pi_2(B)$. In fact $\frac{1}{\sqrt{\pi}} \approx 0.159 > 0.125 = \frac{1}{\sqrt{e}}$. The SVP-algorithm of Prop.1 can use fast BKZ for preprocessing and works even for $\|b_1\| \gg 2\lambda_1$ – see the attack on $\gamma$-unique SVP – whereas [HS07] requires quasy-HKZ-reduction for preprocessing. This eduction already guarantees $\|b_1\| \leq 2\lambda_1$ and performs the main SVP work during preprocessing. Our SVP time bound $n^{n/8+o(n)}$ only assumes $\|b_1\| \leq n^{o(1)} \sqrt{e/\pi} \lambda_1$.

**Theorem 5.4** Given a lattice basis $B \in \mathbb{Z}^{m \times n}$ satisfying GSA and $\|b_1\| \leq \sqrt{\pi} n^b \lambda_1$ for some $b \geq 0$, New Enum solves SVP and proves to have found a solution in time $2^{O(n)} (n^{2+b rd(L)})^{\frac{n}{8}+1+o(n)}$.

Theorem 5.4 is proven in [S10], it does not assume SA and the vol. heuristic. Recall from remark 4 that $n^{\frac{n}{2}+rd rd(L)} \geq 1$ holds under GSA. For $b = o(1)$ Thm. 5.4 shows the SVP-time bound $n^{n/2+o(n)}$ which beats $n^{n/8+o(n)}$ from Hanrot, Stehle [HS07]. Cor. 1 translates Thm. 1 from SVP to CVP, it shows that the corresponding CVP-algorithm solves many important SVP-problems in simple exponential time $2^{O(n)}$ and linear space.

[HS07] proves the time bound $n^{n/2+o(n)}$ for solving CVP by KANNAN’s CVP-algorithm [Ka87]. Minimizing $\|b\|$ for $b \in L \setminus \{0\}$ and minimizing $\|t - b\|$ for $b \in L$ require nearly the same work if $\|t - \ell\| \approx \lambda_1$. In fact the proof of Theorem 1 yields

**New Enum** with linear pruning solves SVP of $L$ of dim $\ell = n$ by (5.4) in worst case heuristic time $n^{n/8+o(n)}$, New Enum solves SVP much faster. Short vectors are found much faster if available stages with large success rate are always performed first and if stages with very small success rate are cut.

6 Primual-dual reduction

**Definition 6.1** Let $B = QR \in \mathbb{R}^{m \times k}$ be a lattice basis with $R = \text{GNF}(B) = [r_{i,j}]_{1 \leq i,j \leq k}$ with blocks $R_\ell = [r_{i,j}]_{1 \leq i \leq \ell, k \leq j \leq k}$, $\ell = 1, \ldots, h$ of size $k$. Then $B$ is a primal-dual basis if

1. it is LLL-basis with HKZ-bases $R_\ell$, $\ell = 1, \ldots, h$.
2. $\max_{\ell = 1, \ldots, h} \max_{k, \ell = 1, \ldots, h} r_{\ell,k}^2 \leq \alpha r_{\ell,k+1}^2$, for $\ell = 1, \ldots, h - 1$, where $r_{\ell,k}^2$ of GNF($R_\ell$,$T$) is maximized over all $T \in \text{GL}_k(\mathbb{Z})$ for $\alpha = 1/(\delta - \frac{1}{2})$ of LLL-reduction.

**Theorem 6.2** [GHKN06] Every primal-dual basis $B = QR \in \mathbb{R}^{m \times k}$ of lattice $L$ satisfies $\|b\|^2 \leq \gamma_k (\alpha \gamma_{1/k}^2 k \lambda^2 (\det R)^{2/\pi})$.

**Proof.** Def. 6.1 shows for $R = [r_{i,j}]_{1 \leq i,j \leq k}$ and $r_{\ell,k}^2$ of GNF($R_\ell$,$T$) that $\max_{T} r_{\ell,k}^2 \leq \alpha r_{\ell+1,k}^2$. The inverse matrix $U_k = \begin{bmatrix} 1 & \cdot \\ \cdot & \cdot \end{bmatrix} \in \mathbb{Z}^{k \times k}$ yields for the lower triangular matrix $R_\ell^{-1} \in \mathbb{R}^{k \times k}$ the upper triangular matrix $R_\ell^T = U_k R_{\ell+1,k} U_k$, where $R_{\ell}^{-1}$ is the transpose of the matrix $R_{\ell}$. The Hermite inequality $\lambda_1^2(L(R_\ell^T)) \leq \gamma_k \lambda_k^{2/\pi}$ for $L(R_\ell^T)$, $\lambda_k = (\det R_k)^{2}$ and HKZ-reduction of $R_\ell^T$ imply

$$D_{\ell}^{1/k} \leq \gamma_k \max_{T} r_{\ell,k}^2 \leq \gamma_k / \lambda_k^2(L(R_\ell^T)).$$

The HKZ-basis $R_{\ell+1}$ satisfies
The combination of these two inequalities and Def. 6.1, part 2, yields
\[ \frac{\lambda_1^2(\mathcal{L}(\mathbf{R}_{t+1}))}{\lambda(t)} = r_{t+k+1,t+1}^2 \leq \gamma_k \mathcal{D}_{t+1}^{1/k}. \]

For the HKZ-basis \( \mathbf{R}_t \) follows by induction over \( t \) that
\[ \| \mathbf{b}_t \| \leq \gamma_k \mathcal{D}_t^{1/k} \leq \gamma_k (\alpha_1)^{t} \mathcal{D}_{t+1}^{1/k} \text{ für } t = 0, \ldots, h - 1. \]

The \( h \)-th root of the product of these \( h \) inequalities proves the claim, as \( \sum_{t=0}^{h-1} \ell = \frac{h^2}{2} \). \( \square \)

Algorithm 6.3 : for Primal-dual reduction

**INPUT** LLL-basis \( \mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_{n+1}] = \mathbf{Q} \mathbf{R} \in \mathbb{Z}^{m \times k}, n + 1 = hk, \) for \( \alpha = 1/(\delta - \frac{1}{4}) \), with blocks \( \mathbf{R}_1, \ldots, \mathbf{R}_h \subset \mathbf{R} \) of size \( k \), \( [\mathbf{b}_{t-k+1}, \ldots, \mathbf{b}_t] = \mathbf{B}_t = \mathbf{Q}_t \mathbf{R}_t, \ell = 1 \)

1. HKZ-reduce the block \( \mathbf{R}_{t+1} \) to \( \mathbf{R}_t \mathbf{T}_k \) with \( \mathbf{T}_k \in \text{GL}_k(\mathbb{Z}) \), \( \mathbf{B}_{t+1} := \mathbf{B}_{t+1} \mathbf{T}_k \)

2. HKZ-reduce \( \mathbf{R}_t^* \) to \( \mathbf{R}_t^* \mathbf{T}_k \) with \( \mathbf{T}_k \in \text{GL}_k(\mathbb{Z}) \), \( \mathbf{B}_t := \mathbf{B}_t \mathbf{T}_k^* \mathbf{U}_k \mathbf{R}_t \), size-reduce \( \mathbf{B}_t, \mathbf{B}_{t+1} \), compute \( \mathbf{R}_{t+1} = \text{GNF}(\mathbf{B}_t, \mathbf{B}_{t+1}) \), LLL-reduce \( \mathbf{R}_{t+1} \) with \( \delta, \alpha \) to \( \mathbf{R}_t \mathbf{T}_k \mathbf{U}_k \)

3. **IF** step 2 exchanged the columns \( k \) and \( k + 1 \) of \( \mathbf{R}_{t+1} \)

   **THEN** \([\mathbf{B}_t, \mathbf{B}_{t+1}] := [\mathbf{B}_t, \mathbf{B}_{t+1}] \mathbf{T}_{2k} \), size-reduce \( \mathbf{B}_t, \mathbf{B}_{t+1}, \ell =: \max(\ell - 1,1) \)

   **ELSE** \( \ell := \ell + 1 \)

4. **IF** \( \ell < h \) **THEN GO TO** 1

**OUTPUT** primal-dual basis \( \mathbf{B} \)

**Comments on Alg. 6.3.** Step 2 maximizes the last diagonal entry \( r_{k,k} \) of \( \text{GNF}(\mathbf{R}_t, \mathbf{T}) \in \mathbb{R}^{k \times k} \) for \( \mathbf{T} \in \text{GL}_k(\mathbb{Z}) \). The revesal matrix \( \mathbf{U}_k := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathbb{Z}^{k \times k} \) yields for \( \mathbf{R}_t^{-1} \in \mathbb{R}^{k \times k} \), the transpose of \( \mathbf{R}_t \), an upper triangular matrix \( \mathbf{R}_t^* \mathbf{U}_k \mathbf{R}_t^{-1} \), \( \mathbf{R}_t^{-1} \) is the transpose of \( \mathbf{R}_t \). HKZ-reduction of \( \mathbf{R}_t^* \) to \( \mathbf{R}_t^* \mathbf{T}_k \) with \( \mathbf{T}_k \in \text{GL}_k(\mathbb{Z}) \) minimizes the first diagonal entry \( r_{1,1} \) of \( \text{GNF}(\mathbf{R}_t^*, \mathbf{T}_k) \) and maximizes \( 1/r_{1,1} \) the last diagonal entry of \( \text{GNF}(\mathbf{R}_t \mathbf{U}_k \mathbf{T}_k^{-1} \mathbf{U}_k) \). The transformation \( \mathbf{T}_k \) of \( \mathbf{R}_t^* \) yields the transformation \( \mathbf{U}_k \mathbf{T}_k^{-1} \mathbf{U}_k \) for \( \mathbf{R}_t \) and \( \mathbf{B}_t \).

The LLL-reduction of \( \mathbf{R}_{t+1} \) in step 2 either starts by exchanging the columns \( k \) and \( k + 1 \) of \( \mathbf{R}_{t+1} \) or \( \mathbf{R}_{t+1} \) is already LLL-reduced, because \( r_{k,k} \) of \( \mathbf{R}_t \) is maximal and \( r_{k+1,k+1} \) of \( \mathbf{R}_{t+1} \) ist minimal.

The primal-dual output \( \mathbf{B} \) of Alg. 6.3 is of the form \( \mathbf{B} = \mathbf{T}_1 \) for the input \( \mathbf{B} \) of Alg.6.3 and \( \mathbf{T}_1 \in \text{GL}_m(\mathbb{Z}) \).

**Theorem 6.4 :** Alg. 6.3 performs at most \( \frac{\alpha^2 t}{\log_4 \alpha} \) iterations before arriving at \( \mathcal{D}_B \leq 1 \).

**Proof.** We replace the Lovász invariance \( \mathcal{D} \) by the following invariance \( \mathcal{D}_B \) where \( \mathcal{D}_t = (\det \mathbf{R}_t)^2 \):
\[ \mathcal{D}_B := \prod_{t=1}^{h-1} (\mathcal{D}_t/\mathcal{D}_{t+1})^{\alpha^2 t/4-(h/2-\ell)^2}. \]

The exponent \( h^2/4-(h/2-\ell)^2 \) is maximal for \( \ell = h/2 \), is zero for \( \ell = 0 \) and \( \ell = h \) and is symmetric to \( \ell = h/2 \). For the LLL-input basis we have \( \mathcal{D}_t \leq \alpha^2 \mathcal{D}_{t+1} \) and therefore her \( \mathcal{D}_B \)-value \( \mathcal{D}_B^{\text{imp}} \) satisfies \( \mathcal{D}_B^{\text{imp}} \leq \alpha^{rk^2} \) for \( s = \sum_{t=1}^{h-1} \ell^2 = h(h-1)(h-1/2)/3 \) yields
\[ \sum_{t=1}^{h-1} (h/2-\ell)^2 = -h^2(h-1)/4 + s = h(h-1)(h-2)/12 \]
and therefore \( s = (h+1)(h-1)/6 = (h^3-h)/6 \).

Hence we have \( \mathcal{D}_B^{\text{imp}} \leq \alpha^{rk^2(h^3-h)/6} \). An active step 3 changes from \( \mathcal{D}_B \) only the factor
\[ \prod_{t=1}^{h-1} (\mathcal{D}_t/\mathcal{D}_{t+1})^{(h/2-\ell)^2}. \]

Every iteration with \( \mathcal{D}_B^{\text{new}} \leq \delta^2 \mathcal{D}_B^{\text{old}} \) decreases as shown \( \mathcal{D}_B \) to \( \mathcal{D}_B^{\text{new}} \leq \delta^2 \mathcal{D}_B^{\text{old}} \). For the number \#It of iterations until arriving at \( \mathcal{D}_B \leq 1 \) we get from \( s = \sum_{t=1}^{h-1} (h^2/4-(h/2-\ell)^2) = (h^3-h)/6 \)
that

\[ \# \text{It} \leq \frac{1}{2} \log_{1/\beta} D_B^{\text{min}} \leq \frac{1}{2} \log_{1/\beta} \alpha k^3 = \frac{k^3}{2} \log_{1/\beta} \alpha. \]

Part 2 of def. 6.1. has been hightend by GAMA, NGUYEN [GN08] to

\[ 2^+ \max_{R_T} r_{k^{\ell+1},k^{\ell+1}} \leq (1 + \varepsilon) r_{k^{\ell+1},k^{\ell+1}} \text{ for } \ell = 1, \ldots, h - 1, \ 0 < \varepsilon \approx 0. \quad \text{[GN08 slide-reduction]} \]

Let \( R'_\ell := [r_{ij}]_{k^{\ell-k-2} \leq i, j \leq k^{\ell+1}} \in R^{k \times k} \) denote the segment one unit to the right of \( R_\ell \). \( \max_{R'_T} r^2_{k^{\ell+1},k^{\ell+1}} \) marks the maximum over \( r^2_{k^{\ell+1},k^{\ell+1}} \) of \( [r_{ij}] = \text{GNF}(R'_T) \) over all \( T \in \text{GL}_k(\mathbb{Z}) \). \( B'_\ell = [b_{k^{\ell-k-2}}, \ldots, b_{k^{\ell+1}}] \) is the block one unit to the right of \( B_\ell \).

**Definition 6.5** A size-reduced basis \( B = QR \in R^{m \times k} \) is *strong primal-dual* if \( R_1, \ldots, R_h \) are HKZ-bases satisfying \( 2^+ \).

**Theorem 6.6** [GN08]

A strong primal-dual basis \( B = QR \in R^{m \times k} \) with \( 0 \approx \varepsilon > 0 \) in \( 2^+ \) of the lattice \( L \) satisfies

\[ ||b_1|| \leq ((1 + \varepsilon) \gamma_k)^{\frac{1}{2}} \frac{b_{h+2}}{\beta^2} (\det L)^{1/bk}. \]

**Proof.** Hermite showed for an HKZ-basis \( R_\ell \) that \( r^2_{k^{\ell-k+1},k^{\ell-k+1}} \leq \gamma_k^2 D_\ell \). The dual of this inequality shows for \( D'_\ell := (\det R'_T)^2 \) that \( \max_{R'_T} r^2_{k^{\ell+k+1},k^{\ell+k+1}} \geq D'_\ell / \gamma_k^2 \) for \( T \in \text{GL}_k(\mathbb{Z}) \).

For \( \ell = \ell_{\max} \), this shows that \( 2^+ \) implies for all primal-dual basis that

\[ D'_\ell \leq (1 + \varepsilon) 2^{2\ell} k^{2k} r^2_{k^{\ell+k+1},k^{\ell+k+1}}. \quad (6.2) \]

Combination of (6.2) with \( r^2_{k^{\ell-k},k^{\ell-k+1}} \leq \gamma_k D_\ell \) and \( D'_\ell / r^2_{k^{\ell+k+1},k^{\ell+k+1}} = D_\ell / r^2_{k^{\ell-k},k^{\ell-k+1}} \) implies

\[ r_{k^{\ell-k},k^{\ell-k+1}} \leq (1 + \varepsilon) (\gamma_k)^{\frac{1}{2k}} r_{k^{\ell-k+1},k^{\ell-k+1}} \quad \text{for } \ell = \ell_{\max} \text{ and } \ell = h - 1. \quad (6.3) \]

For \( \ell = \ell_{\max} \) we get from (6.2) and \( r^2_{k^{\ell-k},k^{\ell-k+1}} \leq \gamma_k D_\ell \) that

\[ D'_\ell \leq (1 + \varepsilon) 2^{2\ell} k^{2k} r^2_{k^{\ell+k+1},k^{\ell+k+1}} \leq (1 + \varepsilon) 2^{2\ell} k^{2k} D_\ell. \quad (6.3) \]

The combination of the two previous inequalities yields for \( \ell = \ell_{\max} \)

\[ D_\ell \leq (1 + \varepsilon) (\gamma_k)^{\frac{1}{2k}} r^2_{k^{\ell+k+1},k^{\ell+k+1}} \leq ((1 + \varepsilon) \gamma_k)^{\frac{1}{2k}} (\det L)^{1/2h}. \quad (6.5) \]

Therefore we get for \( \ell_{\max} \) and also for all \( \ell = 1, \ldots, h - 1 \) that \( D_\ell \leq ((1 + \varepsilon) \gamma_k)^{\frac{1}{2k}} (\det L)^{1/2h}. \)

For the HKZ-basis \( R_1 \) this implies for \( \ell = 1, \ldots, h \) that

\[ ||b_1||^2 \leq \gamma_k D_1^{1/k} \leq \gamma_k ((1 + \varepsilon) \gamma_k)^{\frac{1}{2k-1}} D_1^{1/k}. \]

The product of these \( h \) inequalities and \( \sum_{\ell=1}^h (\ell - 1) = \frac{(h-1)(h)}{2} \) yields

\[ ||b_1||^2 \leq \gamma_k^h ((1 + \varepsilon) \gamma_k)^{\frac{h(h-1)}{2k}} (\det L)^{1/k}. \]

This proves the claim

\[ ||b_1||^2 \leq \gamma_k ((1 + \varepsilon) \gamma_k)^{\frac{1}{2k}} (\det L)^{1/2h} < ((1 + \varepsilon) \gamma_k)^{\frac{h-1}{2k}} (\det L)^{1/2h}. \]

**Algorithm 6.7** : Accelerated, strong primal-dual reduction

**Input** LLL-basis \( B = QR \in Z^{m \times n}, n = hk, 0 \approx \varepsilon > 0. \)

1. Choose \( \ell, 1 \leq \ell < n \) where \( D_\ell / D_{\ell+1} \) is maximal.
2. HKZ-reduce \( R_\ell \) to \( R_\ell + T \) with \( T \in \text{GL}_k(\mathbb{Z}) \), \( B_{\ell+1} := B_{\ell+1}T \), size-reduce \( B_{\ell+1} \), renew \( R_{\ell+1} \). HKZ-reduce \( (R'_T)^* \) to \( (R'_T)^* + T \) with \( T \in \text{GL}_k(\mathbb{Z}) \),

\[ [r_{k^{\ell-k+1},k^{\ell+1}}]_{2 \leq i,j \leq k^{\ell+1}} := \text{GNF}(R'_T \cdot T^{-1} \cdot U_k) \text{ for the reverse matrix } U_k = [1, 1] \in Z^{k \times k}. \]

3. IF \( r_{k^{\ell-k+1},k^{\ell+1}} > (1 + \varepsilon) r_{k^{\ell-k+1},k^{\ell+1}} \) THEN \( B'_\ell := B'_\ell T^{-1} \cdot U_k \), size-reduce \( B'_\ell \).
Theorem 6.8 Alg. 6.7 performs at most \( \frac{k^2h^3}{24} \log_{1+\varepsilon} \alpha \) iterationen until arriving at \( D_B \leq 1 \).

Proof. An active step 3 implies \( D_{\text{new}}^\ell \leq D_{\text{old}}^\ell / (1 + \varepsilon)^2 \) and thus \( D_{\text{new}}^\ell \leq D_{\text{old}}^\ell / (1 + \varepsilon)^4 \). The input-LLL-basis \( B \) satisfies \( D_{\text{inp}}^\ell \leq \alpha^2h^3 - h6 \). Hence \( \text{It} \leq k^2h^3/24 \log_{1+\varepsilon} \alpha \) until \( D_B \leq 1 \).

[GN08] replaces \((1 + \varepsilon)\) in \( 2^+ \) by \( \sqrt{1 + \varepsilon} \) for all \( \ell \leq h - 1 \); [GN08] nearly proves Theorem 6.6 for slide reduced bases. Since we require \( 2^+ \) only for \( \ell_{\text{max}} \) then Alg. 6.7 for strong primal-dual reduction performs at most 2 HKZ-reductions of dim. \( k \) per iteration and therefore is clearly faster than strong primal-dual reduction of [GN08]. Alg. 6.7 for accelerated, strong primal-dual reduction performs about half as many arithmetic operations as alg. 6.3 for primal-dual reduction.

References


