Blockwise Lattice Basis Reduction Revisited.

Claus Peter Schnorr

Fachbereich Informatik und Mathematik, Universität Frankfurt,
PSF 111932, D-60054 Frankfurt am Main, Germany.
schnorr@cs.uni-frankfurt.de – http: www.mi.informatik.uni-frankfurt.de

4. September 2006, work in progress

Abstract. We compare Schnorr's algorithm for semi block $2k$-reduction of lattice bases with Koy's primal-dual reduction for blocksize $2k$. Koy's algorithm guarantees within the same time bound under known proofs better approximations of the shortest lattice vector. Under reasonable heuristics both algorithms are equally strong and much better than proven in worst-case. We combine primal-dual reduction with Schnorr's random sampling reduction (RSR) to a highly parallel reduction algorithm that is on the average more efficient than previous algorithms. It reduces the approximation factor $\frac{4}{3}^{n/2}$ guaranteed by the LLL-algorithm to $1.025^{n/2}$ using feasible lattice reduction.

1 Introduction

A (lattice) *basis* of rank n consists of n linearly independent real vectors $\mathbf{b}_1, ..., \mathbf{b}_n \in \mathbb{R}^m$ which form the basis matrix $B = [\mathbf{b}_1, ..., \mathbf{b}_n] \in \mathbb{R}^{m\times n}$. The basis B generates the *lattice* $\mathcal{L} = \mathcal{L}(B) = \{B\mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^n\} \subset \mathbb{R}^m$ which is the set of all integer linear combinations of the basis vectors. The goal of lattice reduction is to transform a given basis B into a nice basis BT, $T \in \mathrm{SL}_n(\mathbb{R})$, consisting of short and nearly orthogonal vectors.

Notation. $U_k = \begin{bmatrix} & & 1 \\ & \cdot & \\ 1 & & \end{bmatrix} \in \mathbb{Z}^{k\times k}$, the "$U$-turn"-matrix reverses the order of columns/rows of $B \in \mathbb{R}^{m\times n}$ via $B := BU_n$ / $B := U_m B$,
$B = [\mathbf{b}_1, ..., \mathbf{b}_n] \in \mathbb{R}^{m\times n}$, $B^t \in \mathbf{R}^{n\times m}$ is the transpose of $B \in \mathbb{R}^{m\times n}$,
$R^{-t} = (R^{-1})^t = (R^t)^{-1}$ is the inverse transpose of $R \in \mathbb{R}^{n\times n}$,
$d_i = \det([\mathbf{b}_1, ..., \mathbf{b}_i]^t[\mathbf{b}_1, ..., \mathbf{b}_i])$, $d_0 = 1$, $\mathcal{D}_\ell = d_{k\ell}/d_{k\ell-k}$ for given k, B,
$\det \mathcal{L}(B) = \det(B^tB)^{1/2} = d_n^{1/2}$, the ℓ_2-length $\|\mathbf{b}\| = |\mathbf{b}^t\mathbf{b}|^{1/2}$ of $\mathbf{b} \in \mathbb{R}^m$,
$\pi_i : \mathbb{R}^m \to \mathrm{span}(\mathbf{b}_1, ..., \mathbf{b}_{i-1})^\perp$ is the orthogonal projection,
$\lambda_1(\mathcal{L}) =$ the minimal length of the nonzero vectors of the lattice $\mathcal{L}$,
$\gamma_k = \max \lambda_1^2(\mathcal{L}(B))/\det \mathcal{L}(B)^{2/k}$ over all bases $B \in \mathbb{R}^{m\times n}$ of rank k,
$Q \in \mathbb{R}^{m\times n}$ is *isometric* if $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^t\mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$,
$\mathrm{SL}_n(\mathbb{Z}) = \{T \in \mathbb{Z}^{n\times n} \mid \det T = 1\}$, $I_n \in \mathbb{Z}^{n\times n}$ denotes the unit matrix.

The *QR-decomposition* $B = QR$, $R = [\mathbf{r}_1, \ldots, \mathbf{r}_n] = [r_{i,j}] \in \mathbb{R}^{n\times n}$ of a basis $B \in \mathbb{R}^{m\times n}$ consists of an isometric matrix $Q \in \mathbb{R}^{m\times n}$ and an upper-triangular matrix $R \in \mathbb{R}^{n\times n}$ with positive diagonal entries. The QR-decomposition is unique, R is preserved under isometric transforms of B. We call R the *geometric normal form* (GNF) of B, $R = \mathrm{GNF}(B)$. We describe basis reduction in terms of the GNF R. The recent literature on lattice reduction refers to the Gram-Schmidt coefficients $\mu_{j,i} = r_{i,j}/r_{i,i}$, where $r_{i,i} = \|\pi_i(\mathbf{b}_i)\|$ is the length of the vector $\pi_i(\mathbf{b}_i)$. We see from $[\mathbf{b}_1, ..., \mathbf{b}_n] = QR$ that $\|\mathbf{b}_i\|^2 = \sum_{j=1}^{i} r_{j,i}^2$, $\|\mathbf{b}_1\| = r_{1,1}$.

Standard reductions. A lattice basis $B = QR \in \mathbb{R}^{m\times n}$ is *size-reduced* if

$$|r_{i,j}| \le \tfrac{1}{2} r_{i,i} \text{ for all } j > i.$$

$B = QR$ is LLL-*reduced* [LLL82] for $\delta \in (\frac{1}{4}, 1]$ if B is size-reduced and

$$\delta r_{i,i}^2 \le r_{i,i+1}^2 + r_{i+1,i+1}^2 \quad \text{for } i = 1, ..., n-1.$$

Such LLL-bases satisfy $r_{i,i}^2 \le \alpha\, r_{i+1,i+1}^2$ for $\alpha := 1/(\delta - \frac{1}{4})$. Hence

Theorem 1. [LLL82] *An LLL-basis $B \in \mathbb{R}^{m\times n}$ of lattice $\mathcal{L}$ satisfies*

1. $\|\mathbf{b}_1\|^2 \le \alpha^{\frac{n-1}{2}} (\det \mathcal{L})^{2/n}$, **2.** $\|\mathbf{b}_1\|^2 \le \alpha^{n-1} \lambda_1^2$.

The basis $B = QR$ is HKZ-*reduced* (we call it an HKZ-*basis*) if B is size-reduced, and each coefficient $r_{i,i}$ of the GNF R is minimal for all bases of the given lattice that coincide with B in the first $i-1$ vectors.

HKZ-and LLL-reduction are preserved under isometries, i.e., $B = QR$ is HKZ/LLL-reduced iff the GNF R is HKZ/LLL-reduced. HKZ-reduction is due to Hermite [He1850] and Korkine-Zolotareff [KZ1873], LLL-reduction is due to Lenstra, Lenstra, Lovász [LLL82]. The LLL-algorithm transforms a given basis B into an LLL-basis BT, $T \in \mathrm{SL}_n(\mathbb{Z})$. It runs in $O(n^3 m \log_{1/\delta} M_0)$ arithmetic steps using integers of bit length $O(n \log_2 M_0)$, where M_0 denotes $\max\{\|\mathbf{b}_1\|, ..., \|\mathbf{b}_n\|\}$ for the input basis $\mathbf{b}_1, ..., \mathbf{b}_n$. . The LLL-type SLLL-reduction of [S06] runs in $O(n^2 m \log_2 n \log_{1/\delta} M_0)$ arithmetic steps using integers of bit length $2n + \log_2 M_0$. HKZ-reduction runs in $O(mn^{n/2+o(n)})$ arithmetic steps [K87], [S87]. For background and applications of lattice reduction see [MG02].

Survey, background and perspectives of our results. We compare feasible basis reduction algorithms that replace α in Theorem 1 by smaller constants. The approximation factor $\alpha^{\frac{n-1}{2}}$ for $\|\mathbf{b}_1\|/\lambda_1$ has been improved to $2^{O((n \log\log n)^2/\log n)}$ in [S87] and combined with [AKS01] to

$2^{O(n \log\log n/\log n)}$. In this paper we focus on reductions of α achievable in feasible lattice reduction time for practical dimensions. Some reductions are proven by heuristics to be feasible on the average.

Throughout the paper let $\delta \approx 1$ so that $\alpha \approx 4/3$. LLL-bases approximate λ_1 up to a factor $\alpha^{\frac{n-1}{2}} \lesssim 1.155^n$. They approximate λ_1 much better for lattices of *high density* where $\lambda_1^2 \approx \gamma_n (\det \mathcal{L})^{2/n}$, namely up to a factor $\lesssim \alpha^{\frac{n-1}{4}}/\sqrt{\gamma_n} \lesssim 1.075^n$. Moreover, [NS06] reports that α decreases on average to about $1.02^4 \approx 1.08$ for the random lattices of [NS06].

The constant α can be further decreased within polynomial reduction time by blockwise basis reduction. We compare semi block $2k$-reduction [S87] and Koy's primal-dual reduction [K04] with blocksize $2k$. Both algorithms perform HKZ-reductions in dimension $2k$ and have similar polynomial time bounds. They are feasible for $2k \leq 50$.

Semi block $2k$-reduction (**Alg. 1**) replaces α in Theorem 1 by $(\beta_k/\delta)^{1/k}$ for a constant β_k (Theorem 2) that satisfies $k/12 < \beta_k < (1+\frac{k}{2})^{2\ln 2 - 1/k}$ [GHKN06]. Primal-dual reduction (**Alg. 2**) replaces α by $(\alpha\gamma_{2k}^2)^{1/2k}$ (Theorem 3). Since $\gamma_{2k} = \Theta(k)$ the second bound outperforms the first, unless β_k is close to its lower bound $k/12$. Primal-dual reduction for blocks of length 48 replaces α in Theorem 1 within feasible reduction time by $(\alpha\gamma_{48}^2)^{1/48} \approx 1.084$. Both algorithms are equally powerful in approximating λ_1 under the worst-case **GSA**-heuristic of [S03]. They perform under **GSA** much better than proven in worst case without heuristics.

Section 4 uses some basis reduction algorithms that are effcient on average but not proven polynomial time. BKZ-reduction of [SE94] runs in practice for small blocksize in $O(1)$-times the LLL-time bound. The LLL with the *deep insertion* step of [SE94] seems to be polynomial time on the average and greatly improves the approximational power of the LLL. Based on experiments [NS06] reports that deep-insertion-LLL decreases α for random lattices on average to $1.012^4 \approx 1.05 \approx \alpha^{1/6}$.

In section 4 we replace HKZ-reduction within primal-dual reduction by *random sampling reduction* (RSR) of [S03], a parallel extension of the deep insertion step. RSR is nearly feasible up to blocksize $k = 80$. The new algorithm, *primal-dual RSR* (**Alg. 3**) replaces under the worst-case **GSA**-heuristics α in Theorem 1 by $(80/11)^{1/80} \approx 1.025$. **Alg. 3** is highly parallel and polynomial time on the average but not proven polynomial time. If the factor 1.025 further decreases from worst to average case this might endanger the NTRU scheme for $N \leq 200$.

Reductions of $\alpha \approx \frac{4}{3}$ under feasible lattice basis reduction.

1. Semi block $2k$-reduction [S87], $k = 24$
   - proven [GHKN06] $(\beta_{24}/\delta)^{1/24} < 1.165$
   - by heuristic, **GSA** $\gamma_{47}^{1/47} \approx 1.034$
2. Primal-dual HKZ-reduction, Koy 2004, $k = 48$
   - proven $(\alpha\gamma_{48}^2)^{1/48} \approx 1.075$
   - by heuristic, **GSA** $\gamma_{48}^{1/47} \approx 1.034$
3. LLL on the average for random lattices
   - experimentally [NS06] 1.08
4. LLL with deep insertion [SE94]
   - experimentally [NS06] $1.012^4 \approx 1.05$
5. Primal-dual RSR, k=80
   - by heuristic, **GSA, RA** 1.025

GSA is a worst case heuristic, the analysis on the average is open.

2 Semi block $2k$-reduction revisited

Semi block $2k$-reduced bases of [S87] satisfy the inequalities of Theorem 1 with α replaced by $(\beta_k/\delta)^{1/k}$, for a lattice constant β_k such that $\lim_{k\to\infty} \beta_k^{1/k} = 1$. The corresponding reduction algorithm performs HKZ-reductions in dimension $2k$. Let k be given and $n = hk$.

Notation. For a basis $B = QR \in \mathbb{R}^{m\times n}$, $R = [r_{i,j}]_{1\le i,j\le n}$ let

$R_\ell := [r_{i,j}]_{k\ell-k<i,j\le k\ell} \in \mathbb{R}^{k\times k}$ for $k\ell \le n$ and

$R_{\ell,\ell+1} = [r_{i,j}]_{k\ell-k<i,j\le k\ell+k} = \begin{bmatrix} R_\ell & R'_\ell \\ O & R_{\ell+1} \end{bmatrix} \in \mathbb{R}^{2k\times 2k}$ for $\ell < h$

denote the diagonal submatrices of the GNF R corresponding to the segments $B_\ell = [\mathbf{b}_{k\ell+1}, ..., \mathbf{b}_{k\ell+k}]$ and $[B_\ell, B_{\ell+1}]$ of B. We denote

$\mathcal{D}_\ell =_{def} (\det R_\ell)^2 = d_{k\ell}/d_{k\ell-k}$, $\mathcal{D} =_{def} \prod_{\ell=1}^{h-1} d_{k\ell} = \prod_{\ell=1}^{h-1} \mathcal{D}_\ell^{h-\ell}$.

The lattice constant β_k. Let $\beta_k =_{def} \max(\det R_1/\det R_2)^{1/k}$ maximized over all HKZ-reduced GNF's $R = R_{1,2} = \begin{bmatrix} R_1 & R'_1 \\ O & R_2 \end{bmatrix} \in \mathbb{R}^{2k\times 2k}$.

Note that $\beta_1 = \max r_{1,1}^2/r_{2,2}^2$ over all GNF's $R = \left[\begin{smallmatrix} r_{1,1} & r_{1,2} \\ 0 & r_{2,2} \end{smallmatrix}\right] \in \mathbb{R}^{2\times 2}$ satisfying $|r_{1,2}| \le r_{1,1}/2$, $r_{1,1}^2 \le r_{1,2}^2 + r_{2,2}^2$ and thus $\beta_1 = \frac{4}{3} = \alpha$ holds for $\delta = 1$. It is shown in [S87] that $\beta_k \le 4k^2$. This bound has been improved to $\beta_k \le (1+\frac{k}{2})^{2\ln 2 + 1/k}$ in [GHKN06].

Definition 1. [S87] *A basis $B = QR \in \mathbb{R}^{m\times n}$, $n = hk$, is* semi block $2k$-reduced *for $\delta \in (\frac{1}{4}, 1]$ and $\alpha = 1/(\delta - \frac{1}{4})$ if the* GNF $R = [r_{i,j}]$ *satisfies*

1. *$R_1, ..., R_h \subset R$ are HKZ-reduced,*
2. *$r_{i,i}^2 \le \alpha\, r_{i+1,i+1}^2$ for $i = k, 2k, ..., (h-1)k$,*
3. *$\delta\, \mathcal{D}_\ell \le \beta_k^k\, \mathcal{D}_{\ell+1}$ for $\ell = 1, ..., h-1$.*

In [S87] α in clause 2 has been set to 2 and δ in clause 3 has been set to $\frac{3}{4}$. For $k = 1$, clause 3 means that $\delta\, r_{\ell,\ell}^2 \le \frac{4}{3} r_{\ell+1,\ell+1}^2$. This holds if B is LLL-reduced for δ. Clause 1 is empty for $k = 1$, and thus LLL-bases for δ are also semi block $2k$-reduced for $k = 1$. Following [S87] we have

Theorem 2. *A semi block $2k$-reduced basis $B = QR \in \mathbb{R}^{m\times n}$ satisfies*

$$\|\mathbf{b}_1\|^2 \le \gamma_k (\beta_k/\delta)^{\frac{n/k-1}{2}} (\det \mathcal{L}(B))^{2/n}.$$

Moreover, $\|\mathbf{b}_1\|^2 \lambda_1^{-2} \le k^{\ln k + o(\ln k)}\, (\beta_k/\delta)^{n/k-2}$ due to [S87, Thm 3.1, Cor. 3.5]. This replaces α in Theorem 1 by $(\beta_k/\delta)^{1/k}(1 + k^{\frac{1}{2}\ln k}/2n)$.

Proof. We have that $\|\mathbf{b}_1\|^2 \le \gamma_k \mathcal{D}_1^{1/k} = \gamma_k (\det R_1)^{2/k}$ since R_1 is HKZ-reduced. Clause 3 of Def. 1 shows $\mathcal{D}_\ell^{1/k} \le (\beta_k/\delta)\mathcal{D}_{\ell+1}^{1/k}$ and yields

$$\|\mathbf{b}_1\|^2 \le \gamma_k \mathcal{D}_1^{1/k} \le \gamma_k (\beta_k/\delta)^{\ell-1} \mathcal{D}_\ell^{1/k} \quad \text{for } \ell = 1, ..., h = n/k.$$

Multiplying these inequalities and taking h-th roots yields the claim. □

Here is a version of the algorithm in [S87] that does not use the unknown constant β_k. For $k = 1$ it essentially coincides with LLL-reduction.

Alg. 1. Algorithm for semi block $2k$-reduction

`INPUT` basis $B = [\mathbf{b}_1, \ldots, \mathbf{b}_n] \in \mathbb{Z}^{m\times n}$, $\delta \in (\frac{1}{4}, 1)$, $n = hk$.
`OUTPUT` semi block $2k$-reduced basis B.

1. HKZ-reduce B_1, LLL-reduce B, compute $R = [r_{i,j}] = \mathrm{GNF}(B)$, $\ell := 1$.
2. HKZ-reduce $R_{\ell+1}$ into $R_{\ell+1}\, T'$ for some $T' \in \mathrm{SL}_k(\mathbb{Z})$, $B_{\ell+1} := B_{\ell+1}\, T'$, swap $\mathbf{b}_{k\ell}$ and $\mathbf{b}_{k\ell+1}$ if $(r_{k\ell,k\ell}^{new})^2 \le \delta\, r_{k\ell,k\ell}^2$ holds for the new GNF $[r_{i,j}^{new}]$, HKZ-reduce $R_{\ell,\ell+1}$ into $R_{\ell,\ell+1} T$ for some $T \in \mathrm{SL}_{2k}(\mathbb{Z})$.

3. Compute $\mathcal{D}_\ell^{new} := (\det R_\ell^{new})^2$ for $\begin{bmatrix} R_\ell^{new} & R_\ell^{'new} \\ O & R_{\ell+1}^{new} \end{bmatrix} := \mathrm{GNF}(R_{\ell,\ell+1}T)$,
   IF $\mathcal{D}_\ell^{new} \le \sqrt{\delta}\,\mathcal{D}_\ell$ THEN $[B_\ell, B_{\ell+1}] := [B_\ell, B_{\ell+1}]\,T$,
   $\ell := \max(\ell-1, 1)$ ELSE $\ell := \ell+1$.
4. IF $\ell < h$ THEN GO TO 2 ELSE terminate.

The transform T of $R_{\ell,\ell+1}$ is transported to the basis B only if this decreases $\mathcal{D}_\ell$ by the factor $\sqrt{\delta}$. The step $B_{\ell+1} := B_{\ell+1}T'$ does not change $\mathcal{D}_\ell$.

Correctness. Induction over the rounds of the algorithm shows that the basis $\mathbf{b}_1, ..., \mathbf{b}_{k\ell}$ is always semi block $2k$-reduced for the current ℓ. We show that clauses 2, 3 of Def. 1 hold, clause 1 obviously holds.

Clause 2 : If $r_{k\ell,k\ell}^2 > \alpha\, r_{k\ell+1,k\ell+1}^2$ then swapping $\mathbf{b}_{k\ell}$, $\mathbf{b}_{k\ell+1}$ yields

$$(r_{k\ell,k\ell}^{new})^2 \le \tfrac{1}{4} r_{k\ell,k\ell}^2 + r_{k\ell+1,k\ell+1}^2 < (\tfrac{1}{4} + \alpha^{-1}) r_{k\ell,k\ell}^2 = \delta\, r_{k\ell,k\ell}^2,$$

a contradiction since in this case step 2 swaps $\mathbf{b}_{k\ell}$ and $\mathbf{b}_{k\ell+1}$.

Clause 3 : After HKZ-reduction of $R_{\ell,\ell+1}$ we have that $\mathcal{D}_\ell^{new} \le \beta_k^k\, \mathcal{D}_{\ell+1}^{new}$. Since $\mathcal{D}_\ell^{new} > \sqrt{\delta}\,\mathcal{D}_\ell$ and $\mathcal{D}_{\ell+1}^{new} < \sqrt{\delta}\,\mathcal{D}_{\ell+1}$ hold before ℓ gets increased this implies $\sqrt{\delta}\,\mathcal{D}_\ell < \mathcal{D}_\ell^{new} \le \beta_k^k\, \mathcal{D}_{\ell+1}^{new} < \delta^{-1/2}\beta_k^k \mathcal{D}_{\ell+1}$, and thus $\delta\,\mathcal{D}_\ell < \beta_k^k\,\mathcal{D}_{\ell+1}$. □

Lemma 1. *Semi block $2k$-reduction performs at most $h - 1 + 4n(h-1)\log_{1/\delta} M_0$ passes of step 2.*

Proof. Semi block $2k$-reduction iteratively decreases the determinant $\mathcal{D}_\ell$ by HKZ-reduction of $R_{\ell,\ell+1}$. Each pass of step 2 either decreases $\mathcal{D}_\ell$ and $\mathcal{D} = \prod_{\ell=1}^{h-1} \mathcal{D}_\ell^{h-\ell}$ by the factor $\sqrt{\delta}$ or else ℓ is incremented in the subsequent step 3. Since initially $\mathcal{D} = \prod_{\ell=1}^{h-1} d_{k\ell} \le M_0^{2k\binom{h}{2}} = M_0^{n(h-1)}$ the integer $\mathcal{D}$ can be decreased at most $2n(h-1)\log_{1/\delta} M_0$ times by the factor $\sqrt{\delta}$. Hence there are at most $2n(h-1)\log_{1/\delta} M_0$ passes of step 2 that decrease $\mathcal{D}$ by the factor δ, and at most $h - 1 + 2n(h-1)\log_{1/\delta} M_0$ passes of step 2 that do not change $\mathcal{D}$ but increment ℓ in step 3. □

The proof of [S87, Theorem 3.2] shows that a HKZ-reduction of $R_{\ell,\ell+1}$ performs $O(n^2k + k^4 \log M_0) + (2k)^{k+o(k)}$ arithmetic steps using integers of bit length $O(n \log M_0)$. Following [S87], semi block $2k$-reduction performs $O((n^4 + n^2(2k)^{k+o(k)}) \log_{1/\delta} M_0)$ arithmetic steps.

Step 2 tries to decrease $\mathcal{D}_\ell$ within $R_{\ell,\ell+1} \subset R$. In sections 3, 4 we present alternative solutions that are faster than HKZ-reduction of $R_{\ell,\ell+1}$.

3 Primal-dual reduction.

Koy's primal-dual reduction [K04] decreases $\mathcal{D}_\ell$ as follows. It maximizes $r_{k\ell,k\ell}$ over the GNF's of $R_\ell T_\ell$ and minimizes $r_{k\ell+1,k\ell+1}$ over the GNF's of $R_{\ell+1}T_{\ell+1}$ for all T_ℓ, $T_{\ell+1} \in \mathrm{SL}_k(\mathbb{Z})$ and then tries to decrease $r_{k\ell,k\ell}$ and $\mathcal{D}_\ell$ by a subsequent swap of $\mathbf{b}_{k\ell}, \mathbf{b}_{k\ell+1}$. Primal-dual reduction with double blocksize $2k$ replaces the constant β_k/δ in Theorem 2 by $\sqrt{\alpha}\,\gamma_{2k}$ which is better understood than β_k/δ as a function in k, as $\gamma_k = \Theta(k)$.

Dual lattice and dual basis. The *dual* of lattice $\mathcal{L} = \mathcal{L}(QR)$ is the lattice

$$\mathcal{L}^* = \{\mathbf{z} \in \mathrm{span}(\mathcal{L}) \mid \mathbf{z}^t\mathbf{y} \in \mathbb{Z} \text{ for all } \mathbf{y} \in \mathcal{L}\}.$$

Note that $\mathcal{L}^* = \mathcal{L}(QR^{-t})$ because $(QR^{-t})^tQR = R^{-1}Q^tQR = R^{-1}R = I_n$. $QR^{-t} = B^{-t}$ holds for $m = n$.

R^{-t} is a lower triangular matrix and $U_nR^{-t}U_n$ is upper-triangular with positive diagonal entries. Clearly $\mathcal{L}^* = \mathcal{L}(QR^{-t}U_n)$, the basis $QR^{-t}U_n$ has QR-decomposition $QR^{-t}U_n = (QU_n)(U_nR^{-t}U_n)$ because QU_n is isometric and $U_nR^{-t}U_n$ is upper-triangular. $B^* := QR^{-t}U_n$ is the (reversed) *dual basis* of $B = QR$. Note that $(B^*)^* = B$. B^* has the *dual GNF* $R^* := U_nR^{-t}U_n$. The (reversed) dual basis $B^* = [\mathbf{b}_1^*, \ldots, \mathbf{b}_n^*]$ of $B = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$ is characterized by $\langle \mathbf{b}_i^*, \mathbf{b}_{n-j+1}\rangle = \delta_{i,j} = \langle \mathbf{b}_i, \mathbf{b}_{n-j+1}^*\rangle$, where $\delta_{i,j} \in \{0,1\}$ is 1 iff $i = j$.

The diagonal entries of $R = [r_{i,j}]$ and $R^* = [r_{i,j}^*]$ satisfy

$$r_{i,i} = 1/r_{n-i+1,n-i+1}^* \quad \text{for} \quad i = 1, \ldots, n-1. \tag{1}$$

HKZ-reduction of R^* minimizes $r_{1,1}^* = \|\mathbf{b}_1^*\|$ and maximizes $r_{n,n} = 1/r_{1,1}^*$.

Notation. For a basis $B = QR \in \mathbb{R}^{m\times n}$ we let $\bar{r}_{k\ell,k\ell}$ for $k\ell \le n$ denote the maximum of $\widetilde{r}_{k\ell,k\ell}$ over the GNF's $[\widetilde{r}_{i,j}]_{k\ell-k<i,j\le k\ell} = \mathrm{GNF}(R_\ell T)$ for all $T \in \mathrm{SL}_k(\mathbb{Z})$. Shortly, $\bar{r}_{k\ell,k\ell}$ is the maximum of $r_{k\ell,k\ell}$ over the transforms of $R_\ell \subset R$, $\bar{r}_{k\ell,k\ell}$ maximizes $r_{k\ell,k\ell}$ within R_ℓ. .

If $R_\ell^* = U_kR_\ell^{-t}U_k$ is HKZ-reduced then $r_{k\ell,k\ell} = \bar{r}_{k\ell,k\ell}$. We compute $\bar{r}_{k\ell,k\ell}$ by HKZ-reducing R_ℓ^* into R_ℓ^*T, then $[\widetilde{r}_{i,j}] := \mathrm{GNF}(R_\ell U_k T^{-t})$ satisfies $\bar{r}_{k\ell,k\ell} = \widetilde{r}_{k\ell,k\ell}$.

Definition 2. *A basis $B = QR \in \mathbb{R}^{m\times n}$, $n = hk$ is a* primal-dual basis *for k and $\delta \in (\frac{1}{4}, 1]$, $\alpha = 1/(\delta - \frac{1}{4})$ if its* GNF *$R = [r_{i,j}]$ satisfies that*

1. *$R_1, ..., R_h \subset R$ are HKZ-reduced,*
2. *$\bar{r}_{k\ell,k\ell}^2 \le \alpha\, r_{k\ell+1,k\ell+1}^2$ for $\ell = 1, ..., h-1$.*

LLL-bases are primal-dual bases for $k = 1$ since $\bar{r}_{k\ell,k\ell} = r_{k\ell,k\ell}$.

We see from (1) that clause 2 of Def. 2 also holds for the dual B^* of a primal-dual basis B. Therefore, such B^* can be transformed into a primal-dual basis by HKZ-reducing $B^*_\ell = [\mathbf{b}^*_{k\ell+1}, ..., \mathbf{b}^*_{k\ell+k}]$ into $B^*_\ell T_\ell$ for $\ell = 1, ..., h$. Moreover, clauses 2 and 3 of Def. 1 are preserved under duality, they hold for the dual of a semi block $2k$-reduced basis.

Theorem 3. [K04] *A primal-dual basis $B = QR \in \mathbb{R}^{m\times n}$, $n = hk$ of the lattice $\mathcal{L} = \mathcal{L}(B)$ satisfies*

1. $\|\mathbf{b}_1\|^2 \le \gamma_k(\alpha\gamma_k^2)^{\frac{n/k-1}{2}}(\det\mathcal{L})^{2/n}$, **2.** $\|\mathbf{b}_1\|^2 \le (\alpha\gamma_k^2)^{n/k-1}\lambda_1^2$.

Theorem 3 replaces α in Theorem 1 by $(\alpha\gamma_k^2)^{1/k}$. It rephrases for $k = 1, 2$ the bounds of Theorem 1. For $k = 3$ it replaces α in Theorem 1 by $(\alpha\gamma_3^2)^{1/3} = (\alpha 2^{2/3})^{1/3} \lesssim 1.284 < \frac{4}{3}$.

Proof. 1. The maximum $\bar{r}^2_{k\ell,k\ell}$ of $r^2_{k\ell,k\ell}$ over the transforms of R_ℓ satisfies by clause 2 of Def. 2 that

$$\bar{r}^2_{k\ell,k\ell} \le \alpha\, r^2_{k\ell+1,k\ell+1}. \tag{2}$$

Moreover we have $$\mathcal{D}_\ell^{1/k} \le \gamma_k \bar{r}^2_{k\ell,k\ell} = \gamma_k/\lambda_1^2(\mathcal{L}(R^*_\ell)) \tag{3}$$

since $\bar{r}^2_{k\ell,k\ell}$ is computed by HKZ-reduction of R^*_ℓ, and

$$\lambda_1^2(\mathcal{L}(R_{\ell+1})) = r^2_{k\ell+1,k\ell+1} \le \gamma_k \mathcal{D}_{\ell+1}^{1/k} \tag{4}$$

since $R_{\ell+1}$ is HKZ-reduced. Combining these inequalities we get

$$\mathcal{D}_\ell^{1/k} \le \gamma_k \bar{r}^2_{k\ell,k\ell} \le \alpha\gamma_k r^2_{k\ell+1,k\ell+1} \le \alpha\gamma_k^2 \mathcal{D}_{\ell+1}^{1/k}. \tag{5}$$

Since R_1 is HKZ-reduced this yields :

$$\|\mathbf{b}_1\|^2 \le \gamma_k \mathcal{D}_1^{1/k} \le \gamma_k(\alpha\gamma_k^2)^{\ell-1}\mathcal{D}_\ell^{1/k} \quad \text{for } \ell = 1, ..., h.$$

Multiplying these h inequalities and taking h-th roots yields the claim.
2. Note that the inequality (5) also holds for the dual basis B^*, i.e., $(\mathcal{D}^*_\ell)^{1/k} \le \alpha\gamma_k^2(\mathcal{D}^*_{\ell+1})^{1/k}$ holds for $\mathcal{D}^*_\ell = (\det R^*_\ell)^2 = \mathcal{D}^{-1}_{h-\ell}$. Therefore the dual **1*.** of **1.** also holds: $\bar{r}^2_{n,n} \ge \gamma_k^{-1}(\alpha\gamma_k^2)^{-\frac{h-1}{2}}(\det\mathcal{L})^{2/n}$.

1. and **1*.** yield $\|\mathbf{b}_1\|^2 \le \gamma_k^2(\alpha\gamma_k^2)^{h-1}\bar{r}^2_{n,n}$. By clause 2 of Def. 2 we get

$$\|\mathbf{b}_1\|^2 \le \gamma_k^2(\alpha\gamma_k^2)^{\ell-1}\bar{r}^2_{k\ell,k\ell} \le (\alpha\gamma_k^2)^\ell r^2_{k\ell+1,k\ell+1} \quad \text{for} \quad \ell = 0, ..., h-1.$$

Hence $r_{k\ell+1,k\ell+1} \le \lambda_1$ yields the claim. In fact $r_{k\ell+1,k\ell+1} \le \|\pi_{k\ell+1}(\mathbf{b})\| \le \lambda_1$ holds if a shortest lattice vector $\mathbf{b} = \sum_{j=1}^n r_j\mathbf{b}_j \ne \mathbf{0}$ satisfies $k\ell < \mu \le k\ell + k$ for $\mu := \max\{j \mid r_j \ne 0\}$, because $R_{\ell+1}$ is HKZ-reduced. □

Alg. 2. Koy's algorithm for primal-dual reduction
`INPUT` basis $B = QR \in \mathbb{Z}^{m\times n}$, $\delta \in (\frac{1}{4}, 1)$, $n = hk$.
`OUTPUT` primal-dual reduced basis B for k, δ.

1. LLL-reduce B, HKZ-reduce B_1 and compute $R = \mathrm{GNF}(B)$, $\ell := 1$.
2. (*reduce $R_{\ell,\ell+1}$ by primal and dual HKZ-reduction of blocksize k*) HKZ-reduce $R_{\ell+1}$ into $R_{\ell+1}T'$, $B_{\ell+1} := B_{\ell+1}T'$.

   HKZ-reduce $R_\ell^* = U_k R_\ell^{-t} U_k$ into $R_\ell^* T_\ell$. Set $\bar{T} := \begin{bmatrix} U_k T_\ell^{-t} & O \\ O & I_k \end{bmatrix}$.

   LLL-reduce $R_{\ell,\ell+1}\bar{T}$ into $R_{\ell,\ell+1}T$.
3. IF an LLL-swap bridging R_ℓ and $R_{\ell+1}$ occured THEN $[B_\ell, B_{\ell+1}] := [B_\ell, B_{\ell+1}]\,T$, $\ell := \max(\ell-1, 1)$ ELSE $\ell := \ell + 1$.
4. IF $\ell < h$ THEN GO TO 2 ELSE terminate.

Correctness. Induction over the rounds of the algorithm shows that the basis $\mathbf{b}_1, ..., \mathbf{b}_{k\ell}$ is always a primal-dual basis for the current ℓ.

The algorithm deviates from semi block $2k$-reduction only in the steps 2, 3. Step 2 maximizes $r_{k\ell,k\ell}$ within R_ℓ by HKZ-reduction of R_ℓ^* and minimizes $r_{k\ell+1,k\ell+1}$ within $R_{\ell+1}$ by HKZ-reduction of $R_{\ell+1}$, and then LLL-reduces $R_{\ell,\ell+1}$. If no LLL-swap bridging R_ℓ and $R_{\ell+1}$ occured in step 2 then clause 2 of Def. 2 was previously satisfied for ℓ.

Lemma 2. *Primal-dual reduction performs at most $h - 1 + 2n(h-1)\log_{1/\delta} M_0$ passes of step 2.*

Proof. Initially we have $\mathcal{D} = \prod_{\ell=1}^{h-1} d_{\ell k} \le M_0^{n(h-1)}$. Steps 2, 3 either decrease $\mathcal{D}_\ell$ by a factor δ or else increments ℓ. Therefore the proof of Lemma 1 applies. □

The number of passes of step 2 is for both algorithms bounded by $h - 1 + 2n(h-1)\log_{1/\delta} M_0$. The actual number may be smaller but on the average it should be proportional to $h - 1 + 2(h-1)\log_{1/\delta} M_0$.

Alg. 1 and **Alg. 2** (for blocksize $2k$) have by Lemma 1, 2 the same time bound, both algorithms do HKZ-reductions in dimension $2k$.
For double blocksize $2k$ Theorem 3 shows

1. $\|\mathbf{b}_1\|^2 \le \gamma_{2k}(\alpha\gamma_{2k}^2)^{n/4k-1/2}(\det\mathcal{L})^{2/n}$, **2.** $\|\mathbf{b}_1\|^2 \le (\alpha\gamma_{2k}^2)^{n/2k-1}\lambda_1^2$.

These bounds are better than the bounds of Theorem 2 unless β_k is close to the lower bound $\beta_k > k/12$ of [GHKN06] so that $\beta_k/\delta \le \sqrt{\alpha}\,\gamma_{2k}$. But the unknown values β_k can still make semi block $2k$-reduction superior.

The bounds of Theorem 3 are not tight. The bounds of Theorem 3 hold with $\gamma_k^{n/k-1}$ replaced by $\gamma_k(\gamma_k^*)^{n/k-2}$, where we define $\gamma_k^* \le \gamma_k$ as

$$\gamma_k^* := \max[\lambda_1(\mathcal{L})\lambda_1(\mathcal{L}^*)/(\det\mathcal{L}\cdot\det\mathcal{L}^*)^{1/k}]$$

maximized over $\mathcal{L} = \mathcal{L}(B)$ for all bases $B \in \mathbb{R}^{m\times k}$ of rank k.

In particular, primal-dual bases $B = QR \in \mathbb{R}^{m\times n}$, satisfy for $n = hk$

1. $\|\mathbf{b}_1\|^2 \le \sqrt{\alpha}\gamma_k^2(\alpha\gamma_k^{*2})^{\frac{h-2}{2}}(\det\mathcal{L}(B))^{2/n}$, **2.** $\|\mathbf{b}_1\|^2 \le \alpha\gamma_k^2(\alpha\gamma_k^{*2})^{h-2}\lambda_1^2$.

Correctness proof for replacing γ_k^{h-1} by $\gamma_k(\gamma_k^*)^{h-2}$. Theorem 3 follows via (2), (3), (4), (5) from

$$\textstyle\prod_{\ell=1}^{h-1}\lambda_1(\mathcal{L}_\ell^*)\lambda_1(\mathcal{L}_{\ell+1})/[\det\mathcal{L}_\ell^*\det\mathcal{L}_{\ell+1}]^{1/k} \le \gamma_k^{h-1}.$$

By definition of γ_k^* this inequality holds with γ_k^{h-1} replaced by $\gamma_k(\gamma_k^*)^{h-2}$. Both bounds of Theorem 3 and their proof hold with this replacement.

The lattices $\mathcal{L}_k \subset \mathbb{R}^k$ of maximal density are known for dimension $k \le 8$ (see the appendix) and $\mathcal{L}_k$ is unique up to equivalence, i.e., up to isometry and scaling. For $k = 2, 4, 8, 24$ we have that $\mathcal{L}_k^*$ is equivalent to $\mathcal{L}_k$ and thus $\gamma_k = \gamma_k^*$. However $\gamma_k^* < \gamma_k$ holds for $k = 3, 5, 6, 7$ because $\mathcal{L}_k^*$ has lower than maximal density. For instance

$$\lambda_1^2(\mathcal{L}_5^*)/(\det\mathcal{L}_5^*)^{2/5} = 2^{-1/5}\gamma_5 = 2^{-1/5}2^{3/5}.$$

Hence $2^{-1/10}\gamma_5 \le \gamma_5^* < \gamma_5$. If $\mathcal{L}_5$ yields the maximal value γ_5^* then $\gamma_5^* = \sqrt{2} = \gamma_4 = \gamma_4^*$ and γ_k^* would not be strictly increasing in k.

Improving Theorems 2 and 3 by the **GSA***-heuristics.* We associate the quotients $q_i := r_{i+1,i+1}^2/r_{i,i}^2$ with the GNF $R = [r_{i,j}] \in \mathbb{R}^{n\times n}$. The bounds of Theorem 2 and 3 greatly improve under the assumption that nearly all q_i are nearly equal. For simplicity we assume the geometric series assumption (**GSA**) of [S03]:

GSA $$q =_{def} q_1 = q_2 = \cdots = q_{n-1}.$$

GSA is a worst-case property. Bases that do not satisfy **GSA** are easier to reduce, see [S03]. Ajtai's [A01] worst case bases also satisfy **GSA**.

Under **GSA**, α in Theorem 1 can be replaced by q^{-1}. This is because $(\det\mathcal{L})^{2/n}/r_{1,1}^2 = q^{\binom{n}{2}\frac{2}{n}} = q^{n-1}$ holds under **GSA** and thus $r_{1,1}^2 = q^{-n+1}(\det\mathcal{L})^{2/n}$. This replaces α by $1/q$ in part **1.** of Theorem 1. The replacement also holds for part **2.** according to the duality argument in the proof of Theorem 3.

An HKZ-basis R_ℓ satisfies $\lambda_{1,\ell}^2 \le \gamma_k\det(R_\ell)^{2/k}$ by definition of γ_k for $\lambda_{1,\ell} := \lambda_1(\mathcal{L}(R_\ell))$. Under **GSA** this yields

$$1 \le \gamma_k(\det R_\ell)^{\frac{2}{k}}\lambda_{1,\ell}^{-2} = \gamma_k q^{\binom{k}{2}\frac{2}{k}} = \gamma_k q^{k-1},$$

and thus $q \ge \gamma_k^{-1/(k-1)}$. This proves

Corollary 1. *Primal-dual bases of blocksize k satisfy under* **GSA** *the inequalities of Theorem 1 with α replaced by $\gamma_k^{1/(k-1)}$.*

Similarly, $q \geq \gamma_{2k}^{-1/(2k-1)}$ holds for a HKZ-reduced GNF $R = R_{1,2} \in \mathbb{R}^{2k \times 2k}$ under **GSA**. This proves

Corollary 2. *Semi block $2k$-reduced bases satisfy under* **GSA** *the inequalities of Theorem 1 with α replaced by $\gamma_{2k}^{1/(2k-1)}$.*

Primal-dual bases of blocksize $2k$ and semi block $2k$-reduced bases are under **GSA** equally strong, and the upper bounds of Theorems 2, 3 reduce to nearly their square root. This suggests to modify **Alg. 1, 2** to better approximate the **GSA**-property in a way similar to **Alg. 3**.

Practical bounds for $\|\mathbf{b}_1\|^2 \lambda_1^{-2}$. We compare the bounds of Theorems 2 for $2k = 48$ to those of Theorem 3 for double blocksize 48. Note that $\gamma_{24} = 4$, see [CK04]. We assume that $\gamma_{48} \approx 4.9$ corresponding to the densest known lattice packings P_{48p}, P_{48q} in dimension 48, see [CS98, table 1.3]. HKZ-reduction in dimension 48 is nearly feasible. Let $\delta = 0.99$.

Semi block $2k$-reduction for $k = 24$. Using $\beta_{24} \leq 13^{2 \ln 2 + 1/24}$ Theorem 2 proves $\|\mathbf{b}_1\|^2/(\det \mathcal{L})^{2/n} \leq \gamma_{24}(\beta_{24}/\delta)^{n/48-1/2} < \gamma_{24}\, 1.165^{n/2}$. Moreover

$$\|\mathbf{b}_1\|^2/(\det \mathcal{L})^{2/n} \leq \gamma_{48}^{\frac{1}{47}\frac{n-1}{2}}$$

holds under **GSA** replacing α in Theorem 1 by $\gamma_{48}^{1/47} < 1.034$.

Primal-dual bases of blocksize 48 satisfy by Theorem 3

$$\|\mathbf{b}_1\|^2/(\det \mathcal{L})^{2/n} < \gamma_{48}(\alpha\gamma_{48}^2)^{\frac{n/48-1}{2}} \lesssim 1.075^{n/2}/\sqrt{\alpha},$$

replacing α in Theorem 1 by $(\alpha\gamma_{48}^2)^{1/48} \approx 1.075$. Moreover,

$$\|\mathbf{b}_1\|^2/(\det \mathcal{L})^{2/n} \leq \gamma_{48}^{\frac{1}{47}\frac{n-1}{2}}$$

holds under **GSA** replacing α in Theorem 1 by $\gamma_{48}^{1/47} < 1.034$.

Remarks. The bounds under **GSA** indicate that the inequalities of Theorems 2 and 3 are in practice far from being tight. Even the [GHKN06]-lower bound $\beta_k > k/12$ does under **GSA** not limit the power of **Alg. 1** to approximate λ_1 for $k \geq 64$ since β_{64} reduces under **GSA** to $\gamma_{64}^{64/63} < 64/11$. **GSA** is a worst-case heuristics and α further decreases from worst to average case.

The chosen blocksize $k = 48$ has a relatively large γ_k, but the approximational power of **Alg. 1,2** is better when γ_k is relatively low. This makes odd k preferable over even k, in particular since most likely we have that $\gamma_k^* < \gamma_k$. It makes sense to choose k to be fairly distanced to multiples of 24 since γ_k is relatively large if $k = 0 \mod 24$.

4 Primal-dual random sampling reduction.

Primal-dual reduction does not require full HKZ-reductions but merely to replace the first basis vector by a nearly shortest lattice vector. We replace HKZ-reduction by random sampling reduction (RSR) of [S03]. RSR extends the deep insertion step of [SE94] to a highly parallel algorithm. We use RSR in a way to approximate the **GSA**-property. Primal-dual RSR (**ALG. 3**) runs for $k = 80$ in expected feasible time under reasonable heuristics. It reduces α in Theorem 1 to less than $(80/11)^{1/80} \approx 1.025$ which so far is the smallest feasible reduction of α proven under heuristics.

Notation. We associate with a basis $B = QR = [\mathbf{b}_1, ..., \mathbf{b}_n] \in \mathbb{R}^{m \times n}$, $R = [r_{i,j}]_{1 \le i,j \le n}$ the $k \times k$-submatrix $R_{\nu,k} := [r_{i,j}]_{\nu < i,j \le \nu+k} \subset R$ corresponding to $B_{\nu,k} := [\mathbf{b}_{\nu+1}, ..., \mathbf{b}_{\nu+k}] \subset B$. We replace the HKZ-reductions of R_ℓ, R_ℓ^* occuring in **Alg. 2** by RSR of suitable $R_{\nu,k}, R_{\nu,k}^*$.

Alg. 3. Primal-dual RSR
`INPUT` basis $B = QR \in \mathbb{Z}^{m \times n}$, $\delta \in (\frac{1}{4}, 1)$, k.
`OUTPUT` reduced basis B.

1. LLL-reduce B with δ and compute the GNF R of B, $\ell := 0$.
2. IF $\ell = 0 \mod \lfloor n/k \rfloor$ THEN [BKZ-reduce B into BT with δ and blocksize 20, compute the new GNF R, $\ell := \ell + 1$.]
   (*this helps to approximate the* **GSA**-*property.*)
3. *Primal RSR-step.* Randomly select $0 \le \nu \le n - k$ that nearly maximizes $r_{\nu+1,\nu+1}/(\det R_{\nu,k})^{1/k}$. Try to decrease $r_{\nu+1,\nu+1}$ by the factor δ through RSR of $R_{\nu,k} \subset R$, i.e., compute some $T_{\mathbf{a}} = \begin{bmatrix} a_1 & 1 & & \\ \vdots & 0 & \ddots & \\ a_{k-1} & & \ddots & 1 \\ 1 & 0 & \cdots & 0 \end{bmatrix}$ in $\mathrm{GL}_k(\mathbb{Z})$ such that the GNF $\widetilde{R}_{\nu,k} = [\widetilde{r}_{i,j}]$ of $R_{\nu,k} T_{\mathbf{a}}$ satisfies $\widetilde{r}_{\nu+1,\nu+1} \le \delta\, r_{\nu+1,\nu+1}$. For such $T_{\mathbf{a}}$ transform $B_{\nu,k} := B_{\nu,k} T_{\mathbf{a}}$. Recompute $R_{\nu,k}$.
4. *Dual RSR-step.* Randomly select $0 \le \nu \le n - k$ that nearly minimizes $r_{\nu+k,\nu+k}/(\det R_{\nu,k})^{1/k}$. Try to increase $r_{\nu+k,\nu+k}$ by the factor $1/\delta$ through RSR of the dual GNF $R_{\nu,k}^* = U_k R_{\nu,k}^{-t} U_k$, i.e., compute by RSR of $R_{\nu,k}^* = [r_{i,j}^*]$ some $T_{\mathbf{a}} \in \mathrm{GL}_k(\mathbb{Z})$ such that the GNF $\widetilde{R}_{\nu,k}^* = [\widetilde{r}_{i,j}^*]$ of $R_{\nu,k}^* T_{\mathbf{a}}$ satisfies $\widetilde{r}_{\nu+1,\nu+1}^* \le \delta r_{\nu+1,\nu+1}$. (This implies for the GNF $\widetilde{R}_{\nu,k} = [\widetilde{r}_{i,j}]$ of $R_{\nu,k} U_k T_{\mathbf{a}}^{-t}$ that $\widetilde{r}_{\nu+k,\nu+k} \ge r_{\nu+k,\nu+k}/\delta$). For such $T_{\mathbf{a}}$ transform $B_{\nu,k} := B_{\nu,k} U_k T_{\mathbf{a}}^{-t}$. Recompute $R_{\nu,k}$.
5. IF either step 3 or step 4 succeeds, or $\ell \ne 0 \mod \lfloor n/k \rfloor$ THEN GOTO 2 ELSE terminate.

RSR of $R_{\nu,k}$. Let $R_{\nu,k} = [r_{i,j}]_{\nu<i,j\le\nu+k} \subset R = [\mathbf{r}_1, ..., \mathbf{r}_n]$. Enumerate in parallel all vectors $\mathbf{r} := \sum_{j=k/2+1}^{k} a_j \mathbf{r}_{\nu+j} \in \mathcal{L}(R)$ for $(a_{k/2+1}, \ldots, a_k) \in \mathbb{Z}^{k/2}, a_k = 1$ that satisfy $\|\pi_{\nu+k/2+1}(\mathbf{r})\| \le \eta$ for some η having about $(k/11)^{k/4}$ such vectors $\mathbf{r}$. Extend each sum to a short vector $\sum_{j=1}^{k} a_j \mathbf{r}_{\nu+j}$ using $n^2 k^{o(k)}$ bit operations for minimization of $\|\sum_{j=1}^{k} a_j \mathbf{r}_{\nu+j}\|$. We can use the Schnorr-Hörner heuristics [SH95] while full exhaustive search minimization taking $k^{k/4+o(k)} \log_2 M_0$ bit operations is to expensive

RSR extends the SCHNORR-EUCHNER *deep insertion* step [SE94]of depth k. This step coincides with RSR for $a_1, ..., a_{k-1}$ set to zero. RSR tries the zero choice and about $(k/11)^{k/4}$ more instances $(a_1, ..., a_{k-1}) \in \mathbb{Z}^{k-1}$.

Analysis of RSR on $R_{\nu,k}$. Under the assumptions
RA $r_{\nu+j,\nu+j'} \in_R [-\frac{1}{2}, \frac{1}{2}]$ is random for $j' > j$,
GSA $r_{i+1,i+1}/r_{i,i} = q_\nu$ for all i, $\nu < i < \nu + k$,
[S03, Thm 1, Rm. 2] shows that RSR of $R_{\nu,k}$ and $R^*_{\nu,k}$ succeeds in steps 3, 4 as long as $q_\nu < (11/k)^{1/k}$. Primal-dual RSR uses all indices $\nu = 1, ..., n-k$ in a uniform way, this helps to approximate the **GSA**-property.

Theorem 4. [RA, GSA]. *Given a basis $B = QR \in \mathbb{R}^{m\times n}$ of lattice $\mathcal{L} = \mathcal{L}(B)$, primal-dual RSR achieves*

1. $\|\mathbf{b}_1\|^2 \le (k/11)^{\frac{n-1}{2k}} (\det \mathcal{L})^{2/n}$, **2.** $\|\mathbf{b}_1\|^2 \le (k/11)^{\frac{n-1}{k}} \lambda_1^2$.

Proof. Primal-dual RSR terminates with all $q_\nu \ge (11/k)^{1/k}$, proving **1**. The factor $(k/11)^{\frac{n-1}{2k}}$ gets squared by the duality argument in the proof of Theorem 3, this proves **2**. □

Theorem 4 replaces α in Theorem 1 by $(k/11)^{1/k}$, where $(80/11)^{1/80} \approx 1.025$ for $k = 80$.

Primal-dual RSR time bound. RSR succeeds under **RA, GSA** in steps 3 and 4 using $(k/11)^{k/4+o(k)}$ arithm. steps provided that $q_\nu < (11/k)^{1/k}$ [S03, Thm 1, ff.]. For **RA** see [NS06, Fig. 4, 5] (Randomness of $r_{i,i+1}$ is irrelevant, $r_{i,i+1}$ is in practice nearly random in $[-\frac{1}{2}, \frac{1}{2}]$ under the condition that $r_{i,i}^2 \approx r_{i,i+1}^2 + r_{i+1,i+1}^2$, and this improves by deep insertion.) **GSA** is a worst-case assumption, in practice **GSA** is approximately satisfied. [L05] analyses an approximate version of **GSA**.

On the average one round of **Alg. 3** decreases the integer $\mathcal{D}^{(1)} := \prod_{i=1}^{n-1} d_i$ by the factor δ^2. This bounds the average number of rounds by about $\frac{1}{2} n^2 \log_{1/\delta} M_0$ since initially $\mathcal{D}^{(1)} \le M_0^{n(n-1)}$. In worst case however $\mathcal{D}^{(1)}$ can even increase per round and **Alg. 3** must not terminate.

Compairing **Alg. 2** *and* **Alg. 3** *for* $k = 80$. We assume that $\gamma_{80} \approx 4 \cdot 2^{36/80} \approx 5.46$ in view of the densest known lattice $\eta(E_8)$ in dimension 80, see [CS98, table 1.3], we also assume $\gamma_{400} \approx 10.4$.

Under **GSA Alg. 2** reduces α in Theorem 1 for $k = 80$ to $\gamma_{80}^{1/79} < 1.022$.

Primal-dual reduction with full HKZ-reduction is infeasible in dimension $k = 80$ requiring $80^{40+o(1)}$ steps, whereas RSR is nearly feasible.

Primal-dual RSR (**Alg. 3**) reduces by Theorem 4 α in Theorem 1 to $(80/11)^{1/80} < 1.025$ thus achieving $\|\mathbf{b}_1\|/(\det \mathcal{L})^{1/n} < 1.025^{\frac{n-1}{4}}$.

For lattices of high density and $n = 400$, $k = 80$ this yields

$$\|\mathbf{b}_1\|/\lambda_1 < 1.025^{99.75}/\sqrt{\gamma_{400}} \lesssim 3.7.$$

These bounds under **GSA** are worst-case bounds assuming that *all* GNF's R_ν, R_ν^* occuring in the final stages of **Alg. 2, 3** generate lattices of nearly maximal density. This is very unlikely. *Random lattices* have with high probability lower than maximal density. [NS06] reports that the constant $\alpha \approx 4/3$ in Theorem 1 (corresponding to the lattice of maximal density for $n = 2$) reduces for random lattices $R_\nu \in \mathbb{R}^{2\times 2}$ to about $1.02^4 \approx 1.08 < \alpha^{1/3}$. If α further decreases from worst to average case for **Alg. 3**, $k = 80$ from 1.025 to $\leq \sqrt{1.025}$ then the factor 3.7 decreases to ≤ 1.07. This might endanger the NTRU schemes for parameters $N \leq 200$. The secret key is a sufficiently short lattice vector of a lattice of dimension $2N$.

Doing HKZ-reduction via the sieve algorithm of [AKS01] reduces all asymptotic time bounds, but this is unpractical and space consuming.

5 Appendix: Lattices of maximal density.

The lattices $\mathcal{L}_n$ of maximal density for $n = 1, ..., 8$, scaled to $\lambda_1 = 1$ have GNF's $R_n \subset R_8$ consisting of the first n rows and columns of R_8. The coefficients $r_{i,i}$ of $R_8 = [r_{i,j}]$ are minimal for all bases of $\mathcal{L}(R_8)$ that coincide with the first $i-1$ columns of R_8. Size-reduction of R_8 yields HKZ-bases R_n for $n = 1, ..., 8$.

Note that $R_8^t R_8$ has five nonzero diagonals, the coefficients of the main diagonal are 1 and those of the four nonzero side diagonals are $\frac{1}{2}$.

The known Hermite constants γ_n are $\gamma_2 = (\frac{4}{3})^{1/2}$, $\gamma_3 = 2^{1/3}$, $\gamma_4 = 4^{1/2}$, $\gamma_5 = 2^{3/5}$, $\gamma_6 = 23^{-1/6}$, $\gamma_7 = 2^{6/7}$, $\gamma_8 = 2$, $\gamma_{24} = 4$. They correspond to the lattices with basis R_2 – R_8 and to the Leech lattice Λ_{24}.

$$R_8 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{2\cdot 3}} & \frac{1}{2}\sqrt{\frac{3}{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2}\sqrt{\frac{3}{2}} & \frac{1}{\sqrt{2\cdot 3}} & \sqrt{\frac{2}{3}} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{3}} & \frac{1}{2}\frac{1}{\sqrt{3}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

References

[AKS01] *M. Ajtai, R. Kumar, and D. Sivakumar*, A sieve algorithm for the shortest lattice vector problem. In Proc. 33-th STOC, pp. 601-610, ACM, 2001.

[A01] *M. Ajtai*, The worst-case behaviour of Schnorr's algorithm approximating the shortest nonzero vector in a lattice. In Proc. 35-th STOC, pp. 396-406, ACM, 2003.

[CK04] *H. Cohn and A. Kumar*, Optimality and uniqueness of the Leech lattice among lattices. arXiv:math.MG/04 03263v1 16 Mar 2004.

[CS98] *J.H. Conway and N.J.A. Sloane*, Sphere packings, Lattices and Groups. Springer-Verlag, 1998, third edition

[GHKN06] *N. Gama, N. How-Grave-Graham, H. Koy and P. Nguyen*, Rankin's Constant and blockwise lattice reduction. In Proc. CRYPTO'2006, LNCS, Springer-Verlag (to appear) 2006.

[He1850] *C. Hermite*, Extraits de lettres de M. Ch. Hermite à M. Jacobi sur differents objets de la théorie des nombres, deuxième lettre, J. Reine Angewandte Mathematik, **40**, pp. 279–290, 1850.

[Ka87] *R. Kannan*, Minkowski's convex body theorem and integer programming. Mathematics of Operations Research, **12**, pp. 415–440, 1987.

[KZ1873] *A. Korkine und G. Zolotareff*, Sur les formes quadratique, Mathematische Annalen, **6**, pp. 366–389, 1873.

[K04] *H. Koy*, Primal/duale Segment-Reduktion von Gitterbasen, Lecture Universität Franfurt 2000, files from Mai 2004. //www.mi.informatik.uni-frankfurt.de/research/papers.html

[LLL82] *A. K. Lenstra, H. W. Lenstra and L. Lovász*, Factoring polynomials with rational coefficients. *Math. Ann.*, **261**, pp. 515-534, 1982.

[L05] *C. Ludwig*, Practical lattice basis reduction. Dissertation, TU-Darmstadt, December 2005, http://elib.tu-darmstadt.de/diss/000640 and http://deposit.ddb.de/cgi-bin/dokserv?idn=978166493.

[MG02] *D. Micciancio and S. Goldwasser*, Complexity of Lattice Problems, A Cryptographic Perspective. Kluwer Acad. Publ., London, 2002.

[NS06] *P. Nguyen and D. Stehlé*, LLL on the average. In Proc. ANTS-VII, Berlin, 23.-28. July 2006, LNCS, Springer-Verlag, New York, (to appear) 2006.

[S87] *C.P. Schnorr*, A Hierarchy of polynomial time lattice basis reduction algorithms. *Theoret. Comput. Sci.*, **53**, pp. 201–224, 1987.

[S03] *C.P. Schnorr*, Lattice reduction by random sampling and birthday methods. in Proc. STACS 2003, Eds. H. Alt and M. Habib, LNCS 2607, Springer-Verlag, New York, pp. 145–156, 2003. //www.mi.informatik.uni-frankfurt.de/research/papers.html

[S06] *C.P. Schnorr*, Fast LLL-type lattice reduction. Information and Computation, **204**, pp. 1–25, 2006. //www.mi.informatik.uni-frankfurt.de/research/papers.html

[SE94] *C.P. Schnorr and M Euchner*, Lattice basis reduction: improved algorithms and solving subset sum problems. Mathematics of Programming, **66**, pp. 181–189, 1994.

[SH95] *C.P. Schnorr and H.H. Hörner*, Attacking the Chor-Rivest cryptosystem by improved lattice reduction. Proc. Eurocrypt 1995, LNCS 921, Springer-Verlag, pp. 1–12, 1995.