Diophantine Approximation of a Plane

CARSTEN RÖSSNER AND CLAUS P. SCHNORR

Fachbereich Mathematik
Universität Frankfurt
PSTF 11 19 32
60054 Frankfurt/Main, Germany
{roessner,schnorr}@cs.uni-frankfurt.de

May 16, 1997

Abstract

We analyse the problem to approximate a plane by two linearly independent vectors $b_1, b_2 \in \mathbb{Z}^n$. We show that given $x_1 := (x_1^{(1)}, \ldots, x_{n-1}^{(1)}, 1)$, $x_2 := (x_1^{(2)}, \ldots, x_{n-1}^{(2)}, 1) \in \mathbb{R}^n$ our algorithm computes under a reasonable hypothesis for $i = 1, 2, \ldots$ a sequence of linearly independent vectors $b_1^{(i)} := (p_1^{(1,i)}, \ldots, p_{n-1}^{(1,i)}, q^{(1,i)})$, $b_2^{(i)} := (p_1^{(2,i)}, \ldots, p_{n-1}^{(2,i)}, q^{(2,i)}) \in \mathbb{Z}^n$ such that the distance between the planes $x_1 \mathbb{R} + x_2 \mathbb{R}$ and $b_1^{(i)} \mathbb{R} + b_2^{(i)} \mathbb{R}$ is bounded by

$$3^{n/4} \left(\frac{1}{q^{(1,i)\,\frac{1}{n-2}}} + \frac{1}{q^{(2,i)\,\frac{1}{n-2}}} \right)^{1/2} .$$

This generalizes the one–dimensional case where we seek to find for a given vector $x = (x_1, \ldots, x_{n-1}, 1) \in \mathbb{R}^n$ good simultaneous diophantine approximations, i.e., a sequence of vectors $(p_1^{(i)}, \ldots, p_{n-1}^{(i)}, q^{(i)}) \in \mathbb{Z}^n$, $i = 1, 2, \ldots$ such that $\max_{1 \le j \le n} |q^{(i)} x_j - p_j{}^{(i)}|$ becomes arbitrarily small for increasing integers $|q^{(i)}|$. Our result can be extended in a straightforward manner to the r–dimensional case where we seek to approximate an r–dimensional subspace with $1 \le r \le \lfloor \frac{n}{2} \rfloor$.

1 Introduction

We present an algorithm which computes for real $x_1, x_2 \in \mathbb{R}^n$ simultaneous diophantine approximations to the plane spanned by the vectors x_1, x_2. Given linearly independent $x_1, x_2 \in \mathbb{R}^n$ our algorithm constructs a sequence of lattice bases of the lattice $\mathbb{Z}^n$ consisting of vectors that approximate $span(x_1, x_2) := x_1 \mathbb{R} + x_2 \mathbb{R}$. For given $\epsilon > 0$, our algorithm either finds a simultaneous integer relation $m \in \mathbb{Z}^n - 0$ for x_1, x_2 with $< m, x_j >= 0$, $j = 1, 2$ of Euclidean length at most $2^{n/2} \epsilon^{-1}$ or it proves that no simultaneous integer relation exists having length less than ϵ^{-1}. For this the algorithm uses $O(n^4 (n + |\log \epsilon|))$ arithmetic operations on real numbers with exact arithmetic.

For rational input vectors $x_1 := (q_1^{(1)}, \ldots, q_n^{(1)})/q_n^{(1)}$, $x_2 := (q_1^{(2)}, \ldots, q_n^{(2)})/q_n^{(2)}$, with $q_1^{(j)}, \ldots, q_n^{(j)} \in \mathbb{Z}$, $j = 1, 2$ the algorithm has polynomial bit complexity in the input size $\sum_{i=1}^n \lceil \log |q_i^{(1)}| \rceil + \sum_{i=1}^n \lceil \log |q_i^{(2)}| \rceil + |\log \epsilon|$. Our analysis relies on the *dual* lattice basis which we show to consist of very short vectors, see Theorems 1, 2. From this we greatly improve the known bounds for the primary lattice basis and diophantine approximation. The crucial role of the dual basis escaped in all previous studies.

Our algorithm is a generalization of the Stable Continued Fraction Algorithm proposed in [RS96] which in turn is a variant of the HJLS–algorithm of Hastad, Just, Lagarias, Schnorr [HJLS89] incorporating ideas of Just [Ju92] and Rőssner, Schnorr [RS95].

2 Preliminaries

Let $\mathbb{R}^n$ be the n–dimensional real vector space equipped with the ordinary inner product $< .,. >$ and Euclidean length $\|y\| :=< y,y >^{1/2}$. We let $[y_1,\ldots,y_m]$ denote the matrix with column vectors $y_1,\ldots,y_m$ and $\det(y_1,\ldots,y_m)$ denote the volume of the paralleepiped spanned by the vectors $y_1,\ldots,y_m$. Moreover, $\lceil . \rfloor$ is the nearest integer function to a real number r , $\lceil r \rfloor = \lfloor r+0.5 \rfloor$.

A non–zero integral vector $m \in \mathbb{Z}^n$ is called a *simultaneous integer relation* for x_1, x_2 if $< m, x_j >= 0$, $j = 1,2$. We let $\lambda(x_1,x_2)$ denote the length $\|m\|$ of the shortest simultaneous integer relation m for x_1,x_2 , $\lambda(x_1,x_2) = \infty$ if no relation exists.

Throughout this paper, $b_1,\ldots,b_n$ is an ordered basis of the integer lattice $\mathbb{Z}^n$ and its *dual* basis $a_1,\ldots,a_n$ is defined by $[a_1,\ldots,a_n]^\top := [b_1,\ldots,b_n]^{-1}$. Let $x_1,x_2 \in \mathbb{R}^n$ be two linearly independent vectors, set $b_{j-2} := x_j$ for $j = 1,2$. Let π_{x_j} , $j = 1,2$ denote the orthogonal projection onto $(x_j\,\mathbb{R})^\perp$, $j=1,2$ and π_{x_1,x_2} the orthogonal projection onto $(x_1\,\mathbb{R} + x_2\,\mathbb{R})^\perp$. We associate with the basis $b_1,\ldots,b_n$ and the vectors x_1,x_2 the orthogonal projections

$$\begin{aligned}
\pi_{i,x_1,x_2} &: \mathbb{R}^n \longrightarrow span(b_{-1}\,,\ b_0, b_1,\ldots,b_{i-1})^\perp \text{ for } i = -1,\ldots,n \ , \\
\pi_{i,x_j} &: \mathbb{R}^n \longrightarrow span(x_j, b_1,\ldots,b_{i-1})^\perp \text{ for } i = 1,\ldots,n\ ,\ j = 1,2 \text{ and} \\
\pi_i &: \mathbb{R}^n \longrightarrow span(b_1,\ldots,b_{i-1})^\perp \text{ for } i = 1,\ldots,n\,,
\end{aligned}$$

where $span(b_j,\ldots,,b_{i-1})$ denotes the linear space generated by $b_j,\ldots,b_{i-1}$ and $span(b_j,\ldots,b_{i-1})^\perp$ its orthogonal complement in $\mathbb{R}^n$. We abbreviate $\widehat{b}_{i,x_1,x_2} := \pi_{i,x_1,x_2}(b_i)$, $\widehat{b}_{i,x_j} := \pi_{i,x_j}(b_i)$ and $\widehat{b}_i := \pi_i(b_i)$. The vectors $\widehat{b}_{1,x_1,x_2},\ldots,\widehat{b}_{n,x_1,x_2}$ (resp. $\widehat{b}_1,\ldots,\widehat{b}_n$) are pairwise orthogonal. They are called the *Gram–Schmidt orthogonalization* of $x_1,x_2,b_1,\ldots,b_n$ (resp. $b_1,\ldots,b_n$). The *Gram–Schmidt coefficients* $\mu_{i,j}$ of the factorization $[x_1,x_2,b_1,\ldots,b_n] = [x_1,\pi_{x_1}(x_2),\widehat{b}_{1,x_1,x_2},\ldots,\widehat{b}_{n,x_1,x_2}]\ (\mu_{i,j})^\top_{-1\le i,j\le n}$ are defined as $\mu_{i,j} :=< b_i, \widehat{b}_{j,x_1,x_2} > / \|\widehat{b}_{j,x_1,x_2}\|^2$. If $\widehat{b}_{j,x_1,x_2} = 0$, we set $\mu_{i,j} = 0$ for $i \ne j$ and $\mu_{j,j} = 1$. The matrix $(\mu_{i,j})_{-1\le i,j\le n}$ is lower triangular with all diagonal elements 1 . Finally we note that $a_n = \widehat{b}_n / \|\widehat{b}_n\|^2$ since both a_n and $\widehat{b}_n$ are orthogonal to $b_1,\ldots,b_{n-1}$.

The (ordered) projected vectors $\pi_{i,x_1,x_2}(b_i),\ldots,\pi_{i,x_1,x_2}(b_t)$ are *size–reduced* if $|\mu_{k,j}| \le \frac{1}{2}$ holds for $i \le j < k \le t$ and L^3*–reduced* if they are size–reduced and the inequality $\frac{3}{4}\|\pi_{k-1,x_1,x_2}(b_{k-1})\|^2 \ \le \|\pi_{k-1,x_1,x_2}(b_k)\|^2$ holds for $k = i+1,\ldots,t$.

If L^3–reduced the projected vectors $\pi_{i,x_1,x_2}(b_i),\ldots,\pi_{i,x_1,x_2}(b_t)$ satisfy $\|\widehat{b}_{j,x_1,x_2}\|^2 \le 2\,\|\widehat{b}_{j+1,x_1,x_2}\|^2$ for $j = i,\ldots,t-1$.

Models of computation. We distinguish three models of computation for our algorithm.

Exact real arithmetic. For real input $x_1,x_2 \in \mathbb{R}^n$ we use exact arithmetic over real numbers. Our algorithm can use either Gram–Schmidt orthogonalization or Givens Rotation with square roots. The analysis of the HJLS–algortihm applies.

Exact integer arithmetic. For rational input x_1, $x_2 \in \mathbb{Q}^n$ we can use exact arithmetic over the integers. The rational numbers $\mu_{i,j}$, $\|\widehat{b}_{j,x_1,x_2}\|^2$ are represented by their numerator and denominator. This version of our algorithm uses Gram–Schmidt orthogonalization. The analysis of the L^3–algorithm [LLL82] for lattice basis reduction applies.
Floating point arithmetic. For rational input x_1, x_2 we can speed up our algorithm in that we replace the exact arithmetic on the rationals $\mu_{i,j}$, $\|\widehat{b}_{j,x_1,x_2}\|$ by floating point arithmetic. The vectors $x_1, x_2, b_1, \ldots, b_n$, $a_1, \ldots, a_n$ are kept in exact representation. In order to minimize floating point errors we use, instead of the $\mu_{i,j}$, the normalized coefficients $\tau_{i,j} := \mu_{i,j}\,\|\widehat{b}_{j,x_1,x_2}\|$. We call the entities $\tau_{i,j}$ for $-1 \le i, j \le n$ the *orthonormalization* of $x_1, x_2, b_1, \ldots, b_n$. Note that $\tau_{i,i} = \|\widehat{b}_{i,x_1,x_2}\|$. The L^3–property $\frac{3}{4}\,\|\pi_{k-1,x_1,x_2}(b_{k-1})\|^2 \le \|\pi_{k-1,x_1,x_2}(b_k)\|^2$ is expressed by $\frac{3}{4}\,\tau_{k-1,k-1}^2 \le \tau_{k,k}^2 + \tau_{k,k-1}^2$. The $\tau_{i,j}$ are not rational but require square roots, we compute them in floating point arithmetic using Givens Rotation. A complete analysis of the numerical stability of the orthonormalization process can be found in [RS96].

We present our algorithm in its floating point version. From this description the details for the other models of computation are straightforward and left to the reader.

3 The algorithm description

Our algorithm is a variant of the HJLS–algorithm [HJLS89] for finding simultaneous integer relations for real vectors. Our algorithm improves the HJLS–algorithm [HJLS89] towards numerical stability. Given real vectors $x_1, x_2 \in \mathbb{R}^n$ and $\epsilon > 0$, the HJLS–algorithm either finds an integer relation m for x_1, x_2 with $\|m\| \le 2^{n/2-1} \min\{\,\lambda(x_1,x_2)\,,\ \epsilon^{-1}\,\}$ or it proves $\lambda(x_1,x_2) \ge \epsilon^{-1}$. The HJLS–algorithm performs reduction and exchange steps on the linearly dependent system of vectors $x_1, x_2, b_1, \ldots, b_n$ where initially $b_1, \ldots, b_n$ are set to the unit vectors in $\mathbb{R}^n$. We iteratively swap vectors b_{k-1}, b_k, $2 \le k < n$ for which the index $k-1$ maximizes $2^{\sigma(i)}\,\|\widehat{b}_{i,x_1,x_2}\|^2$ where $\sigma(i)$ denotes the number of non–zero orthogonal vectors in $\{\widehat{b}_{1,x_1,x_2}, \ldots, \widehat{b}_{i,x_1,x_2}\}$. Before each swap $b_{k-1} \leftrightarrow b_k$ the vector b_k is size–reduced with respect to its preceding vector b_{k-1}.

The vectors x_1, x_2 remain unchanged and the vectors $b_1, \ldots, b_n$ remain a basis of the lattice $\mathbb{Z}^n$. The HJLS–algorithm uses exact arithmetic on real numbers. Its reduction and exchange steps minimize $\max_{1 \le i \le n} \|\widehat{b}_{i,x_1,x_2}\|$.

The HJLS–algorithm terminates if either $x_1, x_2 \in span(b_1, \ldots, b_{n-1})$, i.e., if a swap $b_{n-1} \leftrightarrow b_n$ results in $\widehat{b}_{n,x_1,x_2} \ne 0$, or if $\max_{1 \le i \le n} \|\widehat{b}_{i,x_1,x_2}\| \le \epsilon$. In the first case, the last vector a_n of the dual basis is a simultaneous integer relation for x_1, x_2. In the latter case, we have $\lambda(x_1,x_2) \ge \epsilon^{-1}$ which follows from

[HJLS89] Proposition 5.2 *Every basis $b_1, \ldots, b_n$ of $\mathbb{Z}^n$ satisfies*

$$\lambda(x_1,x_2) \;\ge\; 1/\max_{1 \le i \le n} \|\widehat{b}_{i,x_1,x_2}\|\ . \tag{1}$$

Our main modifications of the HJLS–algorithm are as follows:

1. Before swapping vectors b_{t-1}, b_t and b_{n-1}, b_n with $\widehat{b}_{t,x_1,x_2} = \widehat{b}_{n,x_1,x_2} = 0$ the projected vectors $\pi_{1,x_1,x_2}(b_1), \ldots, \pi_{1,x_1,x_2}(b_{t-1})$ and $\pi_{t+1,x_1,x_2}(b_{t+1}), \ldots, \pi_{t+1,x_1,x_2}(b_n)$ are L^3–reduced such that $\|\widehat{b}_{t-1,x_1,x_2}\|^2 \le 2\,\|\widehat{b}_{t+1,x_1,x_2}\|^2$.

2. We apply reduction in size, i.e. we reduce b_k so that $|\mu_{k,i}| \le 1/2$ for $i = 1, \ldots, k-1$. Reduction in size has been neglected in [HJLS89] since it is pointless for the exact real arithmetic.

3. In the floating point version orthonormalization of the vectors $x_1, x_2, b_1, \ldots, b_n$ is done by Givens Rotation with a floating point error that is linear in n and $\max_{-1 \le i \le n} \|b_i\|$, see [HT93, Jo93, RS96].

The test on $\tau_{n,n} \neq 0$ actually checks whether $\tau_{n,n} > 2^{-r}$ where r is the number of precision bits of the floating point arithmetic.

In this paper we consider only the approximation of a two–dimensional subspace of the $\mathbb{R}^n$. The algorithm and the analysis given below can be modified in a straightforward manner to the r–dimensional case where we seek to approximate an r–dimensional subspace with $1 \le r \le \lfloor \frac{n}{2} \rfloor$.

Stable diophantine approximation algorithm (SDAA)

INPUT Linearly independent $x_1, x_2 \in \mathbb{R}^n$-0, $\epsilon > 0$.

1. *Initiation.* Let $b_i \in \mathbb{Z}^n$ be the i–th unit vector. Compute the orthonormalization $(\tau_{i,j})_{-1 \le i,j \le n}$ of $x_1, x_2, b_1, \ldots, b_n$ using Givens Rotation (see [RS96]).
$s := 1$, $t := \min\{1 \le j \le n \,:\, \tau_{j,j} = 0\}$.

2. *L^3–reduction of the projected vectors $\pi_{1,x_1,x_2}(b_1), \ldots, \pi_{1,x_1,x_2}(b_{t-1})$ and $\pi_{t+1,x_1,x_2}(b_{t+1}), \ldots, \pi_{t+1,x_1,x_2}(b_{n-1})$ such that* $\|\widehat{b}_{t-1,x_1,x_2}\|^2 \le 2\,\|\widehat{b}_{t+1,x_1,x_2}\|^2$.

WHILE $s \le n-1$ DO:

While there exists k with $1 < k < t$ and $\frac{3}{4}\tau^2_{k-1,k-1} > \tau^2_{k,k} + \tau^2_{k,k-1}$ size–reduce b_k with respect to b_{k-1} by setting $b_k := b_k - \lceil \tau_{k,k-1}/\tau_{k-1,k-1} \rfloor b_{k-1}$, swap b_{k-1} and b_k, and update the orthonormalization using Givens Rotation.
Reduce $b_1, \ldots, b_t$ in size. While $|\tau_{s,s}| \le \epsilon$ increment s to $s+1$.

While $\tau^2_{t-1,t-1} > 2\,\tau^2_{t+1,t+1}$ do

> While there exists k with $t+2 < k < n$ and $\frac{3}{4}\tau^2_{k-1,k-1} > \tau^2_{k,k} + \tau^2_{k,k-1}$ size–reduce b_k with respect to b_{k-1} by setting $b_k := b_k - \lceil \tau_{k,k-1}/\tau_{k-1,k-1} \rfloor b_{k-1}$, swap b_{k-1} and b_k and update the orthonormalization using Givens Rotation.
> Reduce $b_{t+1}, \ldots, b_n$ in size.
> Swap b_{n-1} and b_n, and update the orthonormalization using Givens Rotation.
> od

Swap b_{t-1} and b_t, and update the orthonormalization using Givens Rotation.
While $|\tau_{s,s}| \le \epsilon$ increment s to $s+1$.
OD

Output $(p_1^{(1)}, \ldots, p_{n-1}^{(1)}, q^{(1)}) := b_1$, $(p_1^{(2)}, \ldots, p_{n-1}^{(2)}, q^{(2)}) := b_2$, the next approximation for $span(x_1, x_2)$, see Theorem 4.

3. *Termination.* Compute $[a_1, \ldots, a_n]^\top := [b_1, \ldots, b_n]^{-1}$.
If $\tau_{n,n} > 0$ a simultaneous relation for x_1, x_2 is found. Output the relation a_n for x_1, x_2.
If $s = n$ then $\tau_{i,i} \le \epsilon$ holds for $i = 1, \ldots, n$. Output "$\lambda(x_1, x_2) \ge \epsilon^{-1}$".

Note that the L^3–reduction of the second block $\pi_{t+1,x_1,x_2}(b_{t+1}), \ldots, \pi_{t+1,x_1,x_2}(b_{n-1})$ is embedded in the L^3–reduction of the first block $\pi_{1,x_1,x_2}(b_1), \ldots, \pi_{1,x_1,x_2}(b_{t-1})$. This guarantees the condition $\|\widehat{b}_{t-1,x_1,x_2}\|^2 \le 2\,\|\widehat{b}_{t+1,x_1,x_2}\|^2$ before exchanging both b_{n-1}, b_n and b_{t-1}, b_t.

If $\epsilon = 0$ then SDDA produces a possibly infinite sequence of vectors b_1, b_2 that are good diophantine approximations to $span(x_1, x_2)$.

Correctness properties.
1. Upon termination of step 2 we have $\tau_{i,i} \le \epsilon$ for $i = 1, \ldots, s-1$, $\widehat{b}_{t,x_1,x_2} = 0$ and the projected vectors $\pi_{1,x_1,x_2}(b_1), \ldots, \pi_{1,x_1,x_2}(b_{t-1})$ and $\pi_{t+1,x_1,x_2}(b_1), \ldots, \pi_{t+1,x_1,x_2}(b_{n-1})$ satisfy $\|\widehat{b}_{k,x_1,x_2}\|^2 \le 2\,\|\widehat{b}_{k+1,x_1,x_2}\|^2$, $k \in \{2, \ldots, n\} \setminus \{t\}$ and $\|\widehat{b}_{t-1,x_1,x_2}\|^2 \le 2\,\|\widehat{b}_{t+1,x_1,x_2}\|^2$.
2. Before swapping b_{n-1} and b_n we have $s < n$ (note that $s \neq n$ since $\tau_{n,n} = 0$) and $\tau_{s,s}^{-1} < \epsilon^{-1}$.
Therefore the L^3–reducedness of the projected vectors $\pi_{1,x_1,x_2}(b_1), \ldots, \pi_{1,x_1,x_2}(b_{t-1})$ and $\pi_{t+1,x_1,x_2}(b_1), \ldots, \pi_{t+1,x_1,x_2}(b_{n-1})$ implies that $\|\widehat{b}_{n-1,x_1,x_2}\|^{-1} < 2^{(n-1-t)/2}\,\|\widehat{b}_{t-1,x_1,x_2}\|^{-1} < 2^{(n-2-s)/2}\,\epsilon^{-1}$.

4 Analysis of SDDA in exact real arithmetic

Theorem 1. *Throughout the computation we have* $\|a_n\| \le 2^{(n-1)/2} \min\{\epsilon^{-1},\ \lambda(x_1,x_2)\}$.

Proof. As long as $\widehat{b}_{n-1,x_1,x_2} = 0$ no swap $b_{n-1} \leftrightarrow b_n$ will occur within the computation of SDDA. This leaves the vector $a_n = e_n$ unchanged. So, we may assume $\widehat{b}_{n-1,x_1,x_2} \neq 0$. If the first swap $b_{n-1} \leftrightarrow b_n$, which results in $\widehat{b}_{n-1,x_1,x_2} \neq 0$, also produces $\widehat{b}_{n,x_1,x_2} \neq 0$ we are done since $a_n = e_n$ is a simultaneous relation with length $\|a_n\| = 1$. So, we may further assume $\widehat{b}_{n,x_1,x_2} = 0$. We let $\overline{b}_1, \ldots, \overline{b}_n$, $\overline{a}_1, \ldots, \overline{a}_n$ denote the dual bases before and $b_1, \ldots, b_n$, $a_1, \ldots, a_n$ after an arbitrary swap $b_{n-1} \leftrightarrow b_n$ of SDDA. Let $\overline{\mu}_{i,j}$ be the Gram–Schmidt coefficients and $\widehat{\overline{b}}_{i,x_1,x_2}$ be the orthogonal vectors of $x_1, x_2, \overline{b}_1, \ldots, \overline{b}_n$. We have $\widehat{b}_{n-1,x_1,x_2} = \overline{\mu}_{n,n-1}\,\widehat{\overline{b}}_{n-1,x_1,x_2}$ with $|\overline{\mu}_{n,n-1}| \le \frac{1}{2}$.
From the characterization of a_{n-1} as $< a_{n-1}, b_i >= \delta_{n-1,i}$, which holds throughout the algorithm, we infer that

$$a_{n-1} \;=\; \frac{\widehat{b}_{n-1,x_1,x_2}}{\|\widehat{b}_{n-1,x_1,x_2}\|^2} - \frac{< b_n, \widehat{b}_{n-1,x_1,x_2} >}{\|\widehat{b}_{n-1,x_1,x_2}\|^2}\, a_n \ .$$

Applying this equation to the vectors $\overline{b}_1, \ldots, \overline{b}_n$, $\overline{a}_1, \ldots, \overline{a}_n$ and $|\overline{\mu}_{n,n-1}| \le \frac{1}{2}$ implies the recursion formula

$$\begin{aligned} \|a_n\| \;=\; \|\overline{a}_{n-1}\| &\;=\; \|\widehat{\overline{b}}_{n-1,x_1,x_2}\|^{-1} + |\overline{\mu}_{n,n-1}|\,\|\overline{a}_n\| \\ &\;\le\; \|\widehat{\overline{b}}_{n-1,x_1,x_2}\|^{-1} + \tfrac{1}{2}\,\|\overline{a}_n\| \ . \end{aligned}$$

From the correctness property 2 and inequality (1) we see that

$$2^{(n-2)/2}\,\|\widehat{\overline{b}}_{n-1,x_1,x_2}\| \;\ge\; \max_{1 \le i \le n} 2^{\sigma(i)/2}\,\|\widehat{\overline{b}}_{i,x_1,x_2}\| \;\ge\; 2^{1/2}\,\lambda(x_1,x_2)^{-1} \ ,$$

where $\sigma(i) := |\{1 \le j \le i :\ \widehat{b}_{j,x_1,x_2} \neq 0\}|$. Using $\|\widehat{\overline{b}}_{n-1,x_1,x_2}\|^{-1} \le 2^{(n-3)/2}\,\lambda(x_1,x_2)$ and $\|\widehat{\overline{b}}_{n-1,x_1,x_2}\|^{-1} \le 2^{(n-3)/2}\,\epsilon^{-1}$, which follows from correctness property 2, we can rewrite the recursion formula to

$$\|a_n\| \;\le\; 2^{(n-3)/2}\,\min\{\epsilon^{-1},\ \lambda(x_1,x_2)\} + \tfrac{1}{2}\,\|\overline{a}_n\| .$$

This inequality holds for every exchange $b_{n-1} \leftrightarrow b_n$ of SDDA. Suppose that there are exactly r such exchanges and using that initially $\|a_n\| = 1$ we obtain:

$$\begin{aligned} \|a_n\| &\;\le\; 2^{(n-3)/2}\,\min\{\epsilon^{-1},\ \lambda(x_1,x_2)\} \sum_{j=0}^{r-1} 2^{-j} + 2^{-r} \\ &\;\le\; 2^{(n-1)/2}\,\min\{\epsilon^{-1},\ \lambda(x_1,x_2)\} \ . \end{aligned} \tag{2}$$

Since the inequality $\|a_n\| \le 2^{(n-1)/2} \min\{\epsilon^{-1}, \lambda(x_1,x_2)\}$ holds after any swap $b_{n-1} \leftrightarrow b_n$ it must always hold because a_n does not change between two swaps. $\square$

Theorem 2. *The dual bases $b_1,\ldots,b_n$ and $a_1,\ldots,a_n$, occuring after the L^3–reduction step of* SDDA, *satisfy for $\widehat{b}_{t,x_1,x_2} = 0$*

$$
\begin{aligned}
1.a.\ \|a_i\| &\le 1.5^{n-i}\,(\max_{i \le j < n} \|\widehat{b}_{j,x_1,x_2}\|^{-1} + 2^{(n-1)/2} \min\{\epsilon^{-1}, \lambda(x_1,x_2)\})\ ,\ i = t+1,\ldots,n\ ,\\
1.b.\ \|a_i\| &\le 1.5^{n-i}\,(\max_{\substack{i \le j < n\\ j \ne t}} \|\widehat{b}_{j,x_1,x_2}\|^{-1} + 5 \cdot 2^{(n-1)/2} \min\{\epsilon^{-1}, \lambda(x_1,x_2)\})\ ,\ i = 1,\ldots,t\ ,\\
2.\ \|b_i\| &\le 2^{(n-1)/2} \min\{\epsilon^{-1}, \lambda(x)\}\\
&\quad (4 \cdot 1.5^{n-t+1} \sum_{j=1}^{\min\{i,t-1\}} \prod_{\substack{k=1\\ k \ne j}}^{t-1} \|\widehat{b}_{k,x_1,x_2}\|^{-1} + \sum_{j=t+1}^{i} \prod_{\substack{k=t+1\\ k\ne j}}^{n-1} \|\widehat{b}_{k,x_1,x_2}\|^{-1}) + \sum_{j=1}^{i} \|\widehat{b}_{j,x_1,x_2}\|\ .
\end{aligned}
$$

Proof. 1. Since SDDA did not terminate previously we know that $\widehat{b}_{n,x_1,x_2} = 0$, and thus $\widehat{b}_{j,x_1,x_2} \ne 0$ holds for all indices $j = 1,\ldots,n-1,\ j \ne t$. Let $\mu_{i,j}$ be the Gram–Schmidt coefficients of $x_1, x_2, b_1, \ldots, b_n$ and define the $\nu_{i,j}$ by $(\nu_{i,j})_{1 \le i,j \le n} := (\mu_{i,j})^{-1}_{1 \le i,j \le n}$. We observe that

$$
a_i \;=\; \sum_{\substack{j=i\\ j\ne t}}^{n-1} \nu_{j,i} \frac{\widehat{b}_{j,x_1,x_2}}{\|\widehat{b}_{j,x_1,x_2}\|^2} + \nu_{t,i}\, a_t + \nu_{n,i}\, a_n \tag{3}
$$

holds for $i = 1,\ldots,n$. In fact, since $\mu_{k,t} = 0$ for all $k \ne t$, this formula implies

$$
\begin{aligned}
< a_i, b_k > = &< \sum_{\substack{j=i\\ j\ne t}}^{n-1} \nu_{j,i} \frac{\widehat{b}_{j,x_1,x_2}}{\|\widehat{b}_{j,x_1,x_2}\|^2} + \nu_{t,i}\, a_t + \nu_{n,i}\, a_n\ ,\ \sum_{j=0}^{k} \mu_{k,j}\, \widehat{b}_{j,x_1,x_2} >\\
= &\sum_{\substack{j=1\\ j\ne t,n}}^{n} \nu_{j,i}\, \mu_{k,j} + \nu_{t,i} < a_t, b_k > + \nu_{n,i} < a_n, b_k >\\
= &\ \delta_{i,k} - \nu_{t,k}\, \delta_{t,k} + \nu_{t,i} < a_t, b_k > - \nu_{n,i}\, \mu_{k,n} + \nu_{n,i} < a_n, b_k > = \delta_{i,k}\,.
\end{aligned}
$$

The L^3–reduction step terminates with $|\mu_{i,j}| \le \frac{1}{2}$ for $1 \le j < i \le n$. Hence

$$
|\nu_{i,j}| \le 1.5^{i-j} \quad \text{for} \quad 1 \le j \le i \le n\,.
$$

Now, equation (3) yields for $i = t+1,\ldots,n$

$$
\|a_i\|^2 \;\le\; 1.5^{2(n-i)} \max_{i \le j < n} \|\widehat{b}_{j,x_1,x_2}\|^{-2} + 1.5^{2(n-i)} \|a_n\|^2\,,
$$

and using the correctness property 2 we also have for $i = 1,\ldots,t$

$$
\|a_i\|^2 \;\le\; \sum_{\substack{j=i\\ j \ne t}}^{n-1} 1.5^{2\,(j-i)}\, 2^{j-2} \min\{\epsilon^{-1}, \lambda(x_1,x_2)\}^2 + 1.5^{2(t-i)} \|a_t\|^2 + 1.5^{2(n-i)} \|a_n\|^2\,.
$$

A straightforward induction on the number of decrementations of the index t for which $\widehat{b}_{t,x_1,x_2} = 0$ occurs for the first time shows the following:

$$
\|a_t\|^2 \;\le\; 4 \cdot 1.5^{(n-t+1)}\, 2^{(n-1)/2} \min\{\epsilon^{-1}, \lambda(x_1,x_2)\})\,. \tag{4}
$$

Using the inequality $\|a_n\| \le 2^{(n-1)/2} \min\{\epsilon^{-1}, \lambda(x_1,x_2)\}$ of Theorem 1 proves the first claim:

$$\begin{aligned}
\|a_i\| &\le 1.5^{n-i}\,(\max_{i\le j<n} \|\widehat{b}_{j,x_1,x_2}\|^{-1} + 2^{(n-1)/2} \min\{\epsilon^{-1}, \lambda(x_1,x_2)\})\ ,\ i = t+1,\ldots,n\,,\\
\|a_i\| &\le 1.5^{n-i}\,(\max_{\substack{i\le j<n\\ j\ne t}} \|\widehat{b}_{j,x_1,x_2}\|^{-1} + 5\cdot 2^{(n-1)/2} \min\{\epsilon^{-1}, \lambda(x_1,x_2)\})\ ,\ i = 1,\ldots,t\,.
\end{aligned}$$

2. We rewrite the equations (3) as

$$[a_1,\ldots,a_n] = \left[\frac{\widehat{b}_{1,x_1,x_2}}{\|\widehat{b}_{1,x_1,x_2}\|^2},\ldots,\frac{\widehat{b}_{t-1,x_1,x_2}}{\|\widehat{b}_{t-1,x_1,x_2}\|^2}\,,\,a_t,\,\frac{\widehat{b}_{t+1,x_1,x_2}}{\|\widehat{b}_{t+1,x_1,x_2}\|^2},\ldots,\frac{\widehat{b}_{n-1,x_1,x_2}}{\|\widehat{b}_{n-1,x_1,x_2}\|^2}\,,\,a_n\right](\nu_{i,j})_{1\le i,j\le n}\ .$$

Since the vectors $\widehat{b}_{1,x_1,x_2},\ldots,\widehat{b}_{n-1,x_1,x_2}$ are pairwise orthogonal there exist orthogonal matrices U_l, i.e. $U_l^{-1} = U_l^\top$, $l = 1,2$ such that

$$\left[\frac{\widehat{b}_{1,x_1,x_2}}{\|\widehat{b}_{1,x_1,x_2}\|^2},\ldots,\frac{\widehat{b}_{t-1,x_1,x_2}}{\|\widehat{b}_{t-1,x_1,x_2}\|^2}\,,\,a_t\right] =$$

$$U_1\begin{bmatrix} 1 & & 0 & a'_{t,1}\\ & \ddots & & \vdots\\ 0 & & 1 & a'_{t,t-1}\\ 0 & \ldots & 0 & a'_{t,t}\end{bmatrix}\begin{bmatrix} \|\widehat{b}_{1,x_1,x_2}\|^{-1} & & 0 & \vdots\\ & \ddots & & 0\\ 0 & & \|\widehat{b}_{t-1,x_1,x_2}\|^{-1} & \vdots\\ \ldots & 0 & \ldots & 1\end{bmatrix},$$

with $(a'_{t,1},\ldots,a'_{t,t})^\top := U_1\,a_t$, $a'_{t,t} = \|\widehat{b}_{1,x_1,x_2}\|\cdot\ldots\cdot\|\widehat{b}_{t-1,x_1,x_2}\|$ and

$$\left[\frac{\widehat{b}_{t+1,x_1,x_2}}{\|\widehat{b}_{t+1,x_1,x_2}\|^2},\ldots,\frac{\widehat{b}_{n-1,x_1,x_2}}{\|\widehat{b}_{n-1,x_1,x_2}\|^2}\,,\,a_n\right] =$$

$$U_2\begin{bmatrix} 1 & & 0 & a'_{n,1}\\ & \ddots & & \vdots\\ 0 & & 1 & a'_{n,n-1}\\ 0 & \ldots & 0 & a'_{n,n}\end{bmatrix}\begin{bmatrix} \|\widehat{b}_{t+1,x_1,x_2}\|^{-1} & & 0 & \vdots\\ & \ddots & & 0\\ 0 & & \|\widehat{b}_{n-1,x_1,x_2}\|^{-1} & \vdots\\ \ldots & 0 & \ldots & 1\end{bmatrix},$$

with $(a'_{n,t+1},\ldots,a'_{n,n})^\top := U_2\,a_n$ and $a'_{n,n} = \|\widehat{b}_{1,x_1,x_2}\|\cdot\ldots\cdot\|\widehat{b}_{n-1,x_1,x_2}\|$. From the previous equations and $[b_1,\ldots,b_n]^\top = [a_1,\ldots,a_n]^{-1}$ we see that

$$[b_1,\ldots,b_n] =$$

$$U\begin{bmatrix}
1 & & 0 & \vdots & & & \\
 & \ddots & & 0 & & 0 & \\
0 & & 1 & \vdots & & 0 & \\
\overline{a}_{t,1} & \ldots & \overline{a}_{t,t-1} & \overline{a}_{t,t} & & & \\
 & & & & 1 & 0 & \vdots\\
 & 0 & & & & \ddots & 0\\
 & & 0 & & 0 & 1 & \vdots\\
 & & & & \overline{a}_{n,t+1} & \ldots\ \ \overline{a}_{n,n-1} & \overline{a}_{n,n}
\end{bmatrix}$$

$$\begin{bmatrix} \|\widehat{b}_{1,x_1,x_2}\| & & & & & 0 & & \\ & \ddots & & & & & 0 & \vdots \\ & & \|\widehat{b}_{t-1,x_1,x_2}\| & & & & & \\ & & & 1 & & & & 0 \\ & & & & \|\widehat{b}_{t+1,x_1,x_2}\| & & & \\ 0 & & & & & \ddots & & \vdots \\ & 0 & & & & & \|\widehat{b}_{n-1,x_1,x_2}\| & \\ & \dots & & 0 & & \dots & & 1 \end{bmatrix} (\mu_{i,j})^{\top}_{1\le i,j\le n} \ ,$$

where $\overline{a}_{t,t} := a'^{-1}_{t,t}$, $\overline{a}_{t,i} := -a'_{t,i}/a'_{t,t}$ for $1 < i < t$, $\overline{a}_{t,t} := a'^{-1}_{t,t}$, $\overline{a}_{n,i} := -a'_{n,i}/a'_{n,n}$ for $t < i < n$, respectively, and $U := [U_1 \ U_2]$ is the $n \times n$–blockmatrix consisting of one $n \times t$–block U_1 and one $n \times (n-t)$–block U_2 . Note, that by the orthogonality of U_l , $l = 1, 2$ the matrix U is orthogonal, too. Thus $\|b_i\|$ is the length of the i–th column vector of the cofactor of U in this matrix product. From $\overline{a}_{t,t} = \|\widehat{b}_{1,x_1,x_2}\|^{-1} \cdot \ldots \cdot \|\widehat{b}_{t-1,x_1,x_2}\|^{-1}$ and $\overline{a}_{n,n} = \|\widehat{b}_{t+1,x_1,x_2}\|^{-1} \cdot \ldots \cdot \|\widehat{b}_{n-1,x_1,x_2}\|^{-1}$ and since the matrix $(\mu_{i,j})^{\top}_{1\le i,j\le n}$ is upper triangular with $|\mu_{i,j}| \le 1$ we have for $i = 1, \ldots, n-1$

$$\begin{aligned} \|b_i\|^2 \ &\le \ \|a_t\|^2 \sum_{j=1}^{\min\{i,t-1\}} \prod_{\substack{k=1\\k\ne j}}^{t-1} \|\widehat{b}_{k,x_1,x_2}\|^{-2} \\ &+ \ \|a_n\|^2 \sum_{j=t+1}^{i} \prod_{\substack{k=t+1\\k\ne j}}^{n-1} \|\widehat{b}_{k,x_1,x_2}\|^{-2} + \sum_{\substack{j=1\\j\ne t}}^{i} \|\widehat{b}_{j,x_1,x_2}\|^2 \ . \end{aligned} \tag{5}$$

Now the claimed upper bound of $\|b_i\|$ follows from inequality (4) and from the inequality $\|a_n\| \le 2^{(n-1)/2} \min\{ \lambda(x_1,x_2)\, , \ \epsilon^{-1} \}$ of Theorem 1. $\square$

Lemma 3. *For real input* $x_1 := (x^{(1)}_1, \ldots, x^{(1)}_{n-1}, 1)$, $x_2 := (x^{(2)}_1, \ldots, x^{(2)}_{n-1}, 1) \in \mathbb{R}^n$ *every pair of vectors* $b_1 := (p^{(1)}_1, \ldots, p^{(1)}_{n-1}, q^{(1)})$, $b_2 := (p^{(2)}_1, \ldots, p^{(2)}_{n-1}, q^{(2)}) \in \mathbb{Z}^n$ *occuring after the* L^3*–reduction step of* SDDA *satisfies*

$$\|\widehat{b}_{1,x_1,x_2}\| \ \le \ 1.5^{\frac{n-t+1}{t-1}}\, 2^{\frac{n+4}{2\,(t-1)}+\frac{t-2}{4}}\, \|b_2\|^{-\frac{1}{t-1}} \tag{6}$$

$$\|\widehat{b}_{2,x_1,x_2}\| \ \le \ 1.5^{\frac{n-t+1}{t-1}}\, 2^{\frac{n+5}{2\,(t-1)}+\frac{t-2}{4}+}\, \|b_1\|^{-\frac{1}{t-1}} \ , \tag{7}$$

where $t := \min\{1 \le j \le n \ : \ \widehat{b}_{j,x_1,x_2} = 0\}$.

Proof. If $\widehat{b}_{l,x_1,x_2} = 0$ for either $l = 1$ or $l = 2$ the claim is trivial. Thus, we may assume $t \ge 3$ and $n \ge 4$.

From Theorem 2 and the correctness property 2 we see that

$$\begin{aligned} \|b_1\| \ &\le \ 1.5^{n-t+1}\, 2^{(n+3)/2}\, \epsilon^{-1} \prod_{j=2}^{t-1} \|\widehat{b}_{j,x_1,x_2}\|^{-1} + \|\widehat{b}_{1,x_1,x_2}\| \\ &\le \ 1.5^{n-t+1}\, 2^{(n+4)/2}\, \epsilon^{-1}\, \|\widehat{b}_{2,x_1,x_2}\|^{-(t-2)}\, 2^{(t-2)\,(t-3)/4} \ , \end{aligned}$$

and, analogously,

$$\|b_2\| \ \le \ 1.5^{n-t+1}\, 2^{(n+3)/2}\, \epsilon^{-1} \prod_{j=3}^{t-1} \|\widehat{b}_{j,x_1,x_2}\|^{-1} \left(\|\widehat{b}_{2,x_1,x_2}\|^{-1} + \|\widehat{b}_{1,x_1,x_2}\|^{-1}\right)$$

$$
\begin{aligned}
&+ \quad (\|\widehat{b}_{2,x_1,x_2}\|^2 + \mu_{2,1}^2 \, \|\widehat{b}_{1,x_1,x_2}\|^2)^{1/2} \\
&\leq \quad 1.5^{n-t+1} \, 2^{(n+3)/2} \, \epsilon^{-1} \, \|\widehat{b}_{1,x_1,x_2}\|^{-(t-2)} \, (2^{\frac{(t-3)\,(t-2)}{4}} + 2^{\frac{(t-2)\,(t-1)}{4} - \frac{1}{2}}) + \sqrt{\frac{5}{4}} \\
&\leq \quad 1.5^{n-t+1} \, 2^{(n+4)/2} \, \epsilon^{-1} \, \|\widehat{b}_{1,x_1,x_2}\|^{-(t-2)} \, 2^{\frac{(t-1)\,(t-2)}{4} + 1} \, .
\end{aligned}
$$

A look into the proofs of Theorem 1 and 2 shows that all inequalities, in particular inequality (2), hold with ϵ^{-1} replaced by $(\sqrt{2}\,\|\widehat{b}_{2,x_1,x_2}\|)^{-1}$. Using $\|\widehat{b}_{1,x_1,x_2}\|^2 \leq 2\,\|\widehat{b}_{2,x_1,x_2}\|^2$ inplies the claim. $\square$

For our main result we will use the following

Hypothesis: Let b_1, b_2 be the pair of vectors occuring after the L^3–reduction step of SDDA. *For linearly independent $x_1, x_2 \in \mathbb{R}^n$ the vectors $b_1, \widehat{b}_2$ are independent in the following sense: The probability that the cosine of the angle $\theta :=$ arccos $\frac{<b_2, b_1>}{\|b_2\|\,\|b_1\|}$ is in the interval $[a,b] \subseteq [-1,0]$ is $(b-a)$.*

The hypothesis means that given linearly independent input vectors $x_1, x_2 \in \mathbb{R}^n$ SDDA computes a sequence of vectors $b_1^{(i)}, b_2^{(i)}$, $i = 1, 2, \ldots$ where every $b_1^{(i)}$ and $b_2^{(i)}$ linearly independently converge to $span(x_1, x_2)$.

The hypothesis, although not proven, seems reasonable since the approximation of the plane $span(x_1, x_2)$ amounts to approximating two arbitrary distributed vectors in $span(x_1, x_2)$. We will use the hypothesis in the following form:

$$
Prob_{b_2, b_1 \in L} \left[\frac{< b_2, b_1 >}{\|b_2\|\,\|b_1\|} \; > \; -(1 - \sqrt{\epsilon}) \; | < b_2, b_1 >< 0 \right] \; < \; \sqrt{\epsilon} \, . \tag{8}
$$

We measure the qualtiy of approximation of the plane $span(x_1, x_2)$ by the vectors $b_1^{(i)}, b_2^{(i)}$, $i = 1, 2, \ldots$ in terms of the sine of the angle $\eta^{(i)}$ between the planes $span(b_1^{(i)}, b_2^{(i)})$ and $span(x_1, x_2)$, i.e.,

$$
\sin \eta^{(i)} \; = \; \|\pi_{x_1,x_2}(v^{(i)})\| \; ,
$$

where $v^{(i)} \in span(b_1^{(i)}, b_2^{(i)})$ is a vector orthogonal to the cutting plane $span(b_1^{(i)}, b_2^{(i)}) \cap span(x_1, x_2)$ with unit length 1 .

Theorem 4. *Given $x_1 := (x_1^{(1)}, \ldots, x_{n-1}^{(1)}, 1)$, $x_2 := (x_1^{(2)}, \ldots, x_{n-1}^{(2)}, 1) \in \mathbb{R}^n$ SDDA computes under the above hypothesis a sequence of linearly independent vectors $b_1 := (p_1^{(1,i)}, \ldots, p_{n-1}^{(1,i)}, q^{(1,i)})$, $b_2 := (p_1^{(2,i)}, \ldots, p_{n-1}^{(2,i)}, q^{(2,i)}) \in \mathbb{Z}^n$, $i = 1, 2, \ldots$ such that the following holds with probability $(1 - 2^{1/4} \sqrt{\|\widehat{b}_{2,x_1,x_2}^{(i)}\|})$:*

$$
|\sin \eta^{(i)}| \;\leq\; 3^{n/4} \left(\frac{1}{q^{(1,i)\,\frac{1}{n-2}}} + \frac{1}{q^{(2,i)\,\frac{1}{n-2}}} \right)^{1/2} .
$$

Proof. Let $v^{(i)} =: \lambda_1^{(i)}\, b_1^{(i)} + \lambda_2^{(i)}\, b_2^{(i)}$ with $\lambda_1^{(i)}, \lambda_2^{(i)} \in \mathbb{R}$. By the sine–formula we have

$$
\begin{aligned}
\sin^2 \eta^{(i)} \;&=\; \|\pi_{x_1,x_2}(v^{(i)})\|^2 \\
&=\; \|\lambda_1^{(i)}\, \pi_{x_1,x_2}(b_1^{(i)}) + \lambda_2^{(i)}\, \pi_{x_1,x_2}(b_2^{(i)})\|^2 \\
&=\; (\lambda_1^{(i)} + \mu_{2,1}^{(i)}\, \lambda_2^{(i)})^2 \, \|\widehat{b}_{1,x_1,x_2}\|^2 + \lambda_2^{(i)\,2} \, \|\widehat{b}_{2,x_1,x_2}\|^2 \; ,
\end{aligned}
$$

where $|\mu_{2,1}^{(i)}| := | < b_2^{(i)}, \widehat{b}_{1,x_1,x_2}^{(i)} > | / < \widehat{b}_{1,x_1,x_2}^{(i)}, \widehat{b}_{1,x_1,x_2}^{(i)} > \leq 1/2$ by the size–reduction of $\pi_{x_1,x_2}(b_2^{(i)})$ with respect to $\pi_{x_1,x_2}(b_1^{(i)})$.

From the hypothesis we see that

$$\frac{-< b_2^{(i)}, b_1^{(i)} >}{\|b_2^{(i)}\| \, \|b_1^{(i)}\|} \; < \; (1 - \sqrt{\epsilon})$$

holds with probability greater than $(1-\sqrt{\epsilon})$. Hence for a normalized vector $v^{(i)} =: \lambda_1^{(i)} \, b_1^{(i)} + \lambda_2^{(i)} \, b_2^{(i)}$ and $< b_2^{(i)}, b_1^{(i)} >< 0$ we must have

$$|\lambda_1^{(i)}| \, , \; |\lambda_2^{(i)}| \; < \; \frac{1}{2\sqrt{\epsilon}}$$

with probability greater than $(1 - \sqrt{\epsilon})$.

In the case $< b_1^{(i)}, b_2^{(i)} >> 0$ we clearly have $|\lambda_1^{(i)}| \, , \, |\lambda_2^{(i)}| \, \le 1/\sqrt{2}$. Thus

$$\sin^2 \eta^{(i)} \quad \le \quad \left(\frac{9}{8} \, \|\widehat{b}_{1,x_1,x_2}^{(i)}\|^2 + \frac{1}{2} \, \|\widehat{b}_{2,x_1,x_2}^{(i)}\|^2 \right) \, \frac{1}{\epsilon}$$

holds with probability greater than $(1 - \sqrt{\epsilon})$. Setting $\epsilon := \sqrt{2} \, \|\widehat{b}_{2,x_1,x_2}^{(i)}\|$ implies

$$\sin^2 \eta^{(i)} \quad \le \quad \left(\frac{9}{8} \, \|\widehat{b}_{1,x_1,x_2}^{(i)}\| + \frac{1}{2\sqrt{2}} \, \|\widehat{b}_{2,x_1,x_2}^{(i)}\|^2 \right)$$

with probability greater than $(1 - 2^{1/4} \sqrt{\|\widehat{b}_{2,x_1,x_2}^{(i)}\|})$. From Lemma 3 it follows with probability greater than $(1 - \sqrt{\|\widehat{b}_{2,x_1,x_2}^{(i)}\|})$ that

$$\sin^2 \eta^{(i)} \quad \le \quad 1.5^{\frac{n-t+1}{t-1}} \, 2^{\frac{n+3}{t-1} + \frac{t-2}{4}} \, (\|b_1^{(i)}\|^{-1/(t-1)} + \|b_2^{(i)}\|^{-1/(t-1)}) \; .$$

Since $2 \le t-1 \le n-2$ and either $b_1^{(i)}$ or $b_2^{(i)}$ is linearly independent to the n–th unit vector

$$\begin{aligned} \sin^2 \eta^{(i)} \quad &\le \quad 1.5^{\frac{n-2}{2}} \, 2^{\frac{n+5}{4} + \frac{n-3}{4}} \, (q^{(1,i)^{-1/(n-2)}} + q^{(2,i)^{-1/(n-2)}}) \\ &\le \quad 3^{n/2} \, (q^{(1,i)^{-1/(n-2)}} + q^{(2,i)^{-1/(n-2)}}) \end{aligned}$$

holds with probability greater than $(1 - 2^{1/4} \sqrt{\|\widehat{b}_{2,x_1,x_2}^{(i)}\|})$ for $i = 1, 2, \ldots$. □

Running time. We refer to the models of computation introduced in section 2. Arithmetic operations are $+$, $-$, $\cdot$, $/$, $\lceil \, \rfloor$ (the nearest integer function) and $<$ (comparison). In the floating point version, which has been completely analysed in [RS96], we also use $\sqrt{\ }$ (square root).

Exact real arithmetic. For real input $x \in \mathbb{R}^n$ SDDA performs $O(n^4 \, (n + |\log \epsilon|))$ arithmetic operations on real numbers and $O(n^2 \, (n + |\log \epsilon|))$ many swaps $b_{k-1} \leftrightarrow b_k$ with $2 \le k \le n$. This follows from the analysis of the HJLS–algorithm [HJLS89]. The algorithm either uses Gram–Schmidt orthogonalization via the $\mu_{i,j}$ and $\|\widehat{b}_{j,x_1,x_2}\|^2$ or Givens Rotation with square roots via the $\tau_{i,j}$.

Exact integer arithmetic. For rational $x_j = (q_1^{(j)}, \ldots, q_n^{(j)})/q_n^{(j)} \in \mathbb{Q}^n$, $j = 1, 2$ SDDA performs at most $O(n^4 \, (n + |\log \epsilon|))$ arithmetic operations on integers of bit length $O(n + \max_{\substack{1 \le i \le n \\ j=1,2}} |\log q_i^{(j)}| + |\log \epsilon|)$. Arithmetic steps use the coordinates of the vectors b_i , a_i and the numerators and denominators of the rational numbers $\mu_{i,j}$, $\|\widehat{b}_{j,x_1,x_2}\|^2$. The algorithm uses Gram–Schmidt orthogonalization. The claimed upper bound on the bit length of all integers follows by adjusting the analysis of the L^3–algorithm in [LLL82] to our algorithm.

Computation of a Highly Regular Nearby Plane. SDDA either computes an integer simultaneous relation for x_1, x_2 or proves $\lambda(x_1, x_2) > \epsilon^{-1}$. The algorithm essentially generalizes the Stable Continued Fraction Algorithm (SCFA) of [RS96] which either computes a short integer relation for one input vector x or proves $\lambda(x) > \epsilon^{-1}$. In the latter case SCFA outputs a point $x' \neq x$ and a short integer relation for x'. Moreover, the point x' is *highly regular* in the following sense:

For every $\overline{x} \in \mathbb{R}^n$ with $\|\overline{x} - x\| < \|x' - x\|/2$ there does not exist an integer relation with Euclidean length less than or equal to $1/(2\,\epsilon)$, i.e., $\lambda(\overline{x}) > 1/(2\,\epsilon)$.

It is not clear how to modify SDDA in order to compute points $x_i' \neq x_i$, $i = 1, 2$ and a short simultaneous integer relation for x_1', x_2' such that the plane $span(x_1', x_2')$ is *highly regular* in the following sense: distortions of x_1, x_2 which are (in the Euclidean norm) smaller than $c \max_{i=1,2} \|x_i - x_i'\|$ for some $0 < c < 1/2$ do not significantly destroy the lower bound on the length of the smallest simultaneous integer relation.

Assume, that SDDA stops with the terminal bases $b_1, \ldots, b_n$ and $a_1, \ldots, a_n$ satisfying $\max_{1 \leq i \leq n} \|\widehat{b}_{i,x_1,x_2}\| < \epsilon$. Setting $x_j' := x_j - \pi_n(x_j)$, $j = 1, 2$ yields a short simultaneous integer relation a_n for the nearby points x_1', x_2'. However, these points need not to be highly regular in either sense. (Note that we cannot prove an upper bound on $\max_{\substack{1 \leq i \leq n \\ j=1,2}} \|\widehat{\widehat{b}}_{i,x_j'}\|$ in terms of ϵ and a constant factor alone.)

In order to construct highly regular nearby points x_1', x_2' admitting a short simultaneous integer relation m we choose as a candidate solution $m := \lambda_{n-1}\, a_{n-1} + \lambda_n\, a_n \in span(\widehat{b}_{n-1}, \widehat{b}_n)$ with $\lambda_{n-1}, \lambda_n \in \mathbb{Z}$ and derive the corresponding conditions for 'nearby points–candidates' x_1', x_2' :

Since $a_n = \widehat{b}_n/\|\widehat{b}_n\|^2$ and $a_{n-1} = \widehat{b}_{n-1}/\|\widehat{b}_{n-1}\|^2 - < b_n, \widehat{b}_{n-1} > /\|\widehat{b}_{n-1}\|^2\, b_n$ the condition $< m, x_j' >= 0$, $j = 1, 2$ implies

$$\begin{aligned} 0 &= < \lambda_{n-1}\, a_{n-1} + \lambda_n\, a_n, x_j' > \\ &= \lambda_{n-1} < \frac{\widehat{b}_{n-1}}{\|\widehat{b}_{n-1}\|^2} + \left(\frac{\lambda_n}{\lambda_{n-1}} - \frac{< b_n, \widehat{b}_{n-1} >}{\|\widehat{b}_{n-1}\|^2}\right) \frac{\widehat{b}_n}{\|\widehat{b}_n\|^2}, x_j' > \\ &= \lambda_{n-1} \left[\frac{< \widehat{b}_{n-1}, x_j' >}{\|\widehat{b}_{n-1}\|^2} + \left(\frac{\lambda_n}{\lambda_{n-1}} - \frac{< b_n, \widehat{b}_{n-1} >}{\|\widehat{b}_{n-1}\|^2}\right) \frac{< \widehat{b}_n, x_j' >}{\|\widehat{b}_n\|^2}\right], \; j = 1, 2 \end{aligned}$$

and hence

$$\frac{\lambda_n}{\lambda_{n-1}} = \frac{< b_n, \widehat{b}_{n-1} >}{\|\widehat{b}_{n-1}\|^2} - \frac{< x_j', \widehat{b}_{n-1} > /\|\widehat{b}_{n-1}\|^2}{< x_j', \widehat{b}_n > /\|\widehat{b}_n\|^2}, \; j = 1, 2 \,. \tag{9}$$

Thus we must have

$$\frac{< x_1', \widehat{b}_{n-1} > /\|\widehat{b}_{n-1}\|^2}{< x_1', \widehat{b}_n > /\|\widehat{b}_n\|^2} = \frac{< x_2', \widehat{b}_{n-1} > /\|\widehat{b}_{n-1}\|^2}{< x_2', \widehat{b}_n > /\|\widehat{b}_n\|^2}, \text{ in particular } \pi_{n-1}(x_1') \parallel \pi_{n-1}(x_2') \,.$$

Moreover, it follows that the relation m must satisfy

$$m = \lambda_{n-1} \left(\frac{\widehat{b}_{n-1}}{\|\widehat{b}_{n-1}\|^2} - \frac{< x_j', \widehat{b}_{n-1} > /\|\widehat{b}_{n-1}\|^2}{< x_j', \widehat{b}_n > /\|\widehat{b}_n\|^2} \frac{\widehat{b}_n}{\|\widehat{b}_n\|^2}\right) \in \mathbb{Z}^n, \; j = 1, 2 \,. \tag{10}$$

Now assume that $\|\pi_{n-1,x_1}(x_2)\| < \|\pi_{n-1,x_2}(x_1)\|$. Otherwise we interchange the roles of x_1 and x_2 :

For rational x_1, x_2 we define the nearby points x_1', x_2' by setting

$$\begin{aligned} x_1' &:= x_1 \\ x_2' &:= x_2 - \pi_{n-1,x_1}(x_2) \;=\; x_2 - \pi_{n-1}(x_2) + \frac{< x_2, \pi_{n-1}(x_1) >}{\|\pi_{n-1}(x_1)\|^2}\, \pi_{n-1}(x_1) \;. \end{aligned}$$

For irrational x_1, x_2 we have to take sufficiently close approximations to x_1', x_2'. It is easy to see that from this definition the projections $\pi_{n-1}(x_1')$, $\pi_{n-1}(x_2')$ of x_1', x_2' are parallel, i.e. $\pi_{n-1,x_1'}(x_2') = \pi_{n-1,x_1}(x_2') = \pi_{n-1,x_2'}(x_1) = \pi_{n-1,x_2'}(x_1') = 0$. Moreover, the previously defined vector m is a simultaneous integer relation for x_1', x_2' if

$$\lambda_{n-1}\left(\frac{\widehat{b}_{n-1}}{\|\widehat{b}_{n-1}\|^2} - \frac{< x_1, \widehat{b}_{n-1} > / \|\widehat{b}_{n-1}\|^2}{< x_1, \widehat{b}_n > / \|\widehat{b}_n\|^2}\, \frac{\widehat{b}_n}{\|\widehat{b}_n\|^2}\right) \in \mathbb{Z}^n \;. \tag{11}$$

From the definition of the dual basis vectors we see that

$$< b_n, \widehat{b}_{n-1} > / \|\widehat{b}_{n-1}\|^2 \;=\; - < a_{n-1}, a_n > / \|a_n\|^2 \quad \text{and}$$

$$\widehat{b}_{n-1} / \|\widehat{b}_{n-1}\|^2 \;=\; a_{n-1} - < a_{n-1}, a_n / \|a_n\| > \; a_n / \|a_n\| \;.$$

It follows that

$$\left(\frac{\widehat{b}_{n-1}}{\|\widehat{b}_{n-1}\|^2} - \frac{< x_1, \widehat{b}_{n-1} > / \|\widehat{b}_{n-1}\|^2}{< x_1, \widehat{b}_n > / \|\widehat{b}_n\|^2}\, \frac{\widehat{b}_n}{\|\widehat{b}_n\|^2}\right) \;=\; a_{n-1} - \frac{< x_1, a_{n-1} >}{< x_1, a_n >}\, a_n \;.$$

We thus set

$$\overline{\lambda} \;:=\; \frac{< x_1, a_{n-1} >}{< x_1, a_n >} \quad \text{and compute}$$

a sufficiently close rational approximation $\lambda_n / \lambda_{n-1}$ to $\overline{\lambda}$ via the Euclidean algorithm. Then $m = \lambda_{n-1}\, a_{n-1} + \lambda_n\, a_n$ is the desired simultaneous integer relation for x_1', x_2'.

For rational $x_1 =: (q_1^{(1)}, \ldots, q_{n-1}^{(1)}, q_n^{(1)})^\top / q_n^{(1)}$ with $q_1^{(1)}, \ldots, q_n^{(1)} \in \mathbb{Z}$ the simultaneous integer relation m satisfies

$$\begin{aligned} \|m\| &\leq \|a_{n-1}\|\, |< x_1, a_n >| + \|a_n\|\, |< x_1, a_{n-1} >| \;\leq\; 2\, \|x_1\|\, |q_n^{(1)}|\, \|a_{n-1}\|\, \|a_n\| \\ &\leq 2\left(\sum_{i=1}^{n} q_i^{(1)\,2}\right)^{1/2} \min\{\epsilon^{-1},\ \lambda(x_1, x_2)\}^2\, 2^{(n-1)/2} \cdot 1.5\, \left(2^{(n-3)/2} + 5 \cdot 2^{(n-1)/2}\right) \\ &\leq 9 \cdot 2^n \left(\sum_{i=1}^{n} q_i^{(1)\,2}\right)^{1/2} \min\{\epsilon^{-1},\ \lambda(x_1, x_2)\}^2 \;. \end{aligned}$$

We now prove that the points x_1', x_2' are highly regular:

Proposition 5. *Let $x_1', x_2' \in \mathbb{Q}^n$ be defined as above. Then $\lambda(\overline{x}_1, \overline{x}_2) > 1/(3\,\epsilon)$ holds for all points $\overline{x}_1, \overline{x}_2 \in \mathbb{R}^n$ with*

$$\|x_j - \overline{x}_j\| \;<\; \frac{1}{4}\, \|\pi_{n-1,x_1}(x_2)\| \;=\; \frac{1}{4} \max_{i=1,2} \|x_i - x_i'\| \;.$$

Proof. We use the following

[RS95] Lemma 5(2). *The terminal basis $b_1, \ldots, b_n$ of* SDDA *satisfies the inequalities*

$$\|\widehat{b}_{i,x_j} - \widehat{b}_{i,\overline{x}_j}\| \;\leq\; 2\, \frac{\|\widehat{b}_{i,x_j}\|\, \|\pi_i(x - \overline{x})\|}{\max\{\|\pi_{i+1}(\overline{x}_j)\|\,,\ \|\pi_{i+1}(x_j)\|\}}\ ,\ i = 1, \ldots, n-1\,,\ j = 1, 2\ . \tag{12}$$

This implies for $i = 1, \ldots, n-2$

$$\begin{aligned} \|\widehat{b}_{i,x_1,x_2} - \widehat{b}_{i,\overline{x}_1,x_2}\| &\leq \frac{2\,\|\widehat{b}_{i,x_1,x_2}\|\,\|\pi_{i,x_2}(x_1-\overline{x}_1)\|}{\max\{\|\pi_{i+1,x_2}(\overline{x}_1)\|\,,\,\|\pi_{i+1,x_2}(x_1)\|\}} \\ &\leq \frac{2\,\|\widehat{b}_{i,x_1,x_2}\|\,\|x_1-\overline{x}_1\|}{\|\pi_{n-1,x_2}(x_1)\|} \quad , \quad i=1,\ldots,n-2\,. \end{aligned} \tag{13}$$

and by the same argument

$$\|\pi_{i,x_1}(x_2) - \pi_{i,\overline{x}_1}(x_2)\| \leq 2\,\|\pi_{i,x_1}(x_2)\|\,\frac{\|\pi_i(x_1-\overline{x}_1)\|}{\|\pi_{i,x_2}(x_1)\|} \quad , \quad i=1,\ldots,n-1\,.$$

Since $\|\pi_{n-1,x_2}(x_1)\| \leq \|\pi_{i,x_2}(x_1)\|$, $\|\pi_i(x_1-\overline{x}_1)\| \leq \|x_1-\overline{x}_1\|$ for $i = 1, \ldots, n-1$ we have

$$\|\pi_{i,x_1}(x_2) - \pi_{i,x_1'}(x_2)\| \leq 2\,\|\pi_{i,x_1}(x_2)\|\,\frac{\|x_1-\overline{x}_1\|}{\|\pi_{n-1,x_2}(x_1)\|} \quad , \quad i=1,\ldots,n-1\,.$$

Thus for every $\overline{x}_1$ satisfying $\|x_1-\overline{x}_1\| < \|\pi_{n-1,x_2}(x_1)\|/r_1$ with an $r_1 > 2$

$$0 < \|\pi_{i,\overline{x}_1}(x_2)\| < \frac{r_1+2}{r_1}\,\|\pi_{i,x_1}(x_2)\| \quad , \quad i=1,\ldots,n-1\,.$$

Furthermore

$$0 < \|\widehat{b}_{i,\overline{x}_1,x_2}\| \quad < \quad \frac{r_1+2}{r_1}\,\|\widehat{b}_{i,x_1,x_2}\| < \frac{r_1+2}{r_1}\,\epsilon\,. \tag{14}$$

By an analogous analysis we obtain for $i = 1, \ldots, n-2$

$$\begin{aligned} \|\widehat{b}_{i,\overline{x}_1,x_2} - \widehat{b}_{i,\overline{x}_1,\overline{x}_2}\| &\leq \frac{2\,\|\widehat{b}_{i,\overline{x}_1,x_2}\|\,\|\pi_{i,\overline{x}_1}(x_2-\overline{x}_2)\|}{\max\{\|\pi_{i+1,\overline{x}_1}(x_2)\|\,,\,\|\pi_{i+1,\overline{x}_1}(\overline{x}_2)\|\}} \\ &\leq \frac{2\,r_1}{r_1-2}\,\frac{\|\widehat{b}_{i,\overline{x}_1,x_2}\|\,\|x_2-x_2'\|}{\|\pi_{i+1,x_1}(x_2)\|} \;\leq\; \frac{2\,r_1}{r_1-2}\,\frac{\|\widehat{b}_{i,\overline{x}_1,x_2}\|\,\|x_2-\overline{x}_2\|}{\|\pi_{n-1,x_1}(x_2)\|} \\ &\leq 2\,\frac{r_1/r_2}{r_1-2}\,\|\widehat{b}_{i,\overline{x}_1,x_2}\| \;\leq\; 2\,\frac{r_1+2}{r_1-2}\,\frac{\|x_2-\overline{x}_2\|}{\|\pi_{n-1,x_1}(x_2)\|}\,\|\widehat{b}_{i,x_1,x_2}\| \end{aligned}$$

for every $\overline{x}_2$ satisfying $\|x_2-\overline{x}_2\| < \|\pi_{n-1,x_1}(x_2)\|/r_2$ with an $r_2 > \frac{2\,r_1}{r_1-2}$.
Inequality (9), (10) and (11) and the choice of r_2 imply

$$\begin{aligned} \|\widehat{b}_{i,x_1,x_2} - \widehat{b}_{i,\overline{x}_1,\overline{x}_2}\| &\leq \frac{2}{r_1}\,\|\widehat{b}_{i,x_1,x_2}\| + \frac{2}{r_2}\,\frac{r_1+2}{r_1-2}\,\|\widehat{b}_{i,x_1,x_2}\| \\ &\leq \left(\frac{2}{r_1} + \frac{r_1+2}{r_1}\right)\|\widehat{b}_{i,x_1,x_2}\| \;=\; \frac{r_1+4}{r_1}\,\|\widehat{b}_{i,x_1,x_2}\| \end{aligned}$$

and $\|\widehat{b}_{i,x_1,x_2}\| > 0$ for $i = 1, \ldots, n-2$. This yields $\|\widehat{b}_{i,\overline{x}_1,\overline{x}_2}\| < 2\,\frac{r_1+2}{r_1}\,\epsilon$ for $i = 1, \ldots, n-2$ and every pair of vectors $\overline{x}_1$, $\overline{x}_2$ satisfying

$$\begin{aligned} \|x_1-\overline{x}_1\| &< \|\pi_{n-1,x_2}(x_1)\|/r_1 \qquad \text{and} \\ \|x_2-\overline{x}_2\| &< \frac{r_1-2}{2\,r_1}\,\|\pi_{n-1,x_1}(x_2)\| \text{ with an } r_1 > 2\,. \end{aligned}$$

For $r_1 := 4$ it follows from $\|x_1-\overline{x}_1\| < \|\pi_{n-1,x_2}(x_1)\|/4$ and $\|x_2-\overline{x}_2\| < \|\pi_{n-1,x_1}(x_2)\|/4$ that $\|\widehat{b}_{i,\overline{x}_1,\overline{x}_2}\| < 3\,\epsilon$ holds for $i = 1, \ldots, n-2$.
Using Proposition 5.2 of [HJLS89] and $\pi_{n-1,x_1'}(x_2') = 0$ completes the proof of the claim.
□

5 Open Problems

A still open problem and challenging task is to prove Theorem 4 without using the hypothesis. Also replacing the right hand side of the Theorem's inequality with a term $1/(q_1^{(1,i)}\, q_2^{(2,i)})^{\frac{1}{2\,(n-1)}}\, c^n$ for some constant $c > 0$ is desirable with respect to the Dirichlet bound [Di1842].

As mentioned above SDDA can be modified in that it computes a highly regular nearby plane $span(x_1', x_2')$, where the spanning vectors x_1', x_2' admit a simultaneous integer relation.

In order to generalize the result of [RS96] (Theorem 1) we would have to prove an upper bound of the Euclidean length of the relation which is linear in $\min\{\epsilon^{-1}, \lambda(x_1, x_2)\}$ <u>and</u> $\lambda(x_1', x_2')$. It is even still open whether such a bound holds for rational input vectors x_1, x_2 .

We would be rather interested in helpful comments and/or joint work in order to improve this preliminary draft to a comprehensive paper.

References

[Di1842] G.L. Dirichlet: Verallgemeinerung eines Satzes aus der Lehre von den Kettenbrűchen nebst einigen Anwendungen auf die Theorie der Zahlen, Bericht űber die zur Bekanntmachung geigneten Verhandlungen der Kőniglich Preussischen Akademie der Wissenschaften zu Berlin (1842), pp. 93–95.

[FB92] H.R.P. Ferguson and D.H. Bailey: A Polynomial Time, Numerically Stable Integer Relation Algorithm. RNR Technical Report RNR-91-032, NASA Ames Research Center, Moffett Field, CA (1992).

[HT93] C. Heckler and L. Thiele: A Parallel Lattice Basis Reduction for Mesh–connected Processor Arrays and Parallel Complexity. Proceedings of the 5th Symposium on Parallel and Distributed Processing, Dallas (1993).

[HJLS89] J. Hastad, B. Just, J.C. Lagarias and C.P. Schnorr: Polynomial Time Algorithms for Finding Integer Relations among Real Numbers. SIAM J. Comput., Vol. 18, No. 5 (1989), pp. 859–881.

[Jo93] A. Joux: A Fast Parallel Lattice Basis Reduction Algorithm. Proceedings of the 2nd Gauss Symposium, Munich (1993).

[Ju92] B. Just: Generalizing the Continued Fraction Algorithm to Arbitrary Dimensions. SIAM J. Comput., Vol. 21, No. 5 (1992), pp. 909–926.

[LLL82] A.K. Lenstra, H.W. Lenstra,Jr. and L. Lovász: Factoring Polynomials with Rational Coefficients. Math. Ann. 21 (1982), pp. 515–534.

[RS95] C. Rőssner and C.P. Schnorr: Computation of Highly Regular Nearby Points. Proceedings of the 3rd Israel Symposium on Theory of Computing and Systems, Tel Aviv (1995).

[RS96] C. Rőssner and C.P. Schnorr: An Optimal, Stable Continued Fraction Algorithm. Proceedings of the 5th Conference on Integer Programming and Combinatorial Optimization, Vancouver, B.C. (1996).