The Complexity of Approximate Optima for Greatest Common Divisor Computations

Carsten Rössner and Jean-Pierre Seifert*

Dept. of Math. Comp. Science
University of Frankfurt
P.O. Box 111932
60054 Frankfurt/Main
Germany
{roessner,seifert}@cs.uni-frankfurt.de

Abstract. We study the approximability of the following **NP**-complete (in their feasibility recognition forms) number theoretic optimization problems:

1. Given n numbers $a_1, \ldots, a_n \in \mathbb{Z}$, find a *minimum* gcd *set for* $a_1, \ldots, a_n$, i.e., a subset $S \subseteq \{a_1, \ldots, a_n\}$ with minimum cardinality satisfying $\gcd(S) = \gcd(a_1, \ldots, a_n)$.
2. Given n numbers $a_1, \ldots, a_n \in \mathbb{Z}$, find a ℓ_∞*-minimum* gcd *multiplier for* $a_1, \ldots, a_n$, i.e., a vector $\mathbf{x} \in \mathbb{Z}^n$ with minimum $\max_{1 \leq i \leq n} |x_i|$ satisfying $\sum_{i=1}^{n} x_i a_i = \gcd(a_1, \ldots, a_n)$.

We present a polynomial-time algorithm which approximates a minimum gcd set for $a_1, \ldots, a_n$ within a factor $1+\ln n$ and prove that this algorithm is best possible in the sense that unless $\mathbf{NP} \subseteq \mathbf{DTIME}(n^{O(\log\log n)})$, there is no polynomial-time algorithm which approximates a minimum gcd set within a factor $(1-o(1))\ln n$.

Concerning the second problem, we prove under the slightly stronger complexity theory assumption, $\mathbf{NP} \not\subseteq \mathbf{DTIME}(n^{\mathrm{poly}(\log n)})$, that there is no polynomial-time algorithm which approximates a ℓ_∞-minimum gcd multiplier within a factor $2^{\log^{1-\gamma} n}$, where γ is an arbitrary small positive constant.

Complementary to this result, there exists a polynomial-time algorithm, which computes a gcd multiplier $\mathbf{x} \in \mathbb{Z}^n$ for $a_1, \ldots, a_n \in \mathbb{Z}$ with $\|\mathbf{x}\|_\infty \leq 0.5\,\|\mathbf{a}\|_\infty$. In this paper, we also present a simple polynomial-time algorithm which computes a gcd multiplier $\mathbf{x} \in \mathbb{Z}^n$ with Euclidean length $\|\mathbf{x}\| \leq 1.5^n \|\mathbf{a}\| / \gcd(a_1, \ldots, a_n)$.

Our inapproximability results rely on gap-preserving reductions from minimization problems with equal inapproximability ratios. We implicitly use the close connection between the hardness of approximation and the theory of interactive proof systems, particularly the work of [3, 9, 17, 14].

Key Words. Approximation algorithm, computational complexity, gcd, label cover, **NP**-hard, number theoretic problems, probabilistically checkable proofs, 2-prover 1-round interactive proof systems

* Supported by DFG under grant DFG-Leibniz-Programm Schn 143/5-1

1 Introduction

It is widely believed that **NP**-optimization problems cannot be solved efficiently, i.e., in time polynomial in the input length of the problem at hand. However, in many practical applications approximate solutions of **NP**-optimization problems suffice. Thus, there has been done a lot of work in studying the complexity of finding such approximate solutions for **NP**-optimization problems.

On the one hand, it is desirable to have approximation algorithms such that the value of the returned solution is within a small factor to the optimum solution of the problem. Considering minimization problems, the worst-case ratio of the value of the solution returned by the approximation algorithm to the optimum solution is called the *approximation factor* of the approximation algorithm. Considering a minimization problem Π with input instance I we say that the function $opt_{\Pi}(\cdot)$ is approximable within a factor $f(I)$ if there exists a polynomial-time algorithm A such that $A(I) \le f(I)\, opt_{\Pi}(I)$.

On the other hand, we want to guarantee that the constructed approximation algorithms are best possible in that no substantial better approximation factors can be achieved, unless certain complexity theoretical assumptions, e.g., $\mathbf{NP} \not\subseteq \mathbf{DTIME}(n^{\mathrm{poly}(\log n)})$, are wrong.

In this paper we investigate the following **NP**-complete optimization problems:

Minimum GCD Set (MinGCDS)
INSTANCE: n numbers $a_1, \ldots, a_n \in \mathbb{Z}$
SOLUTION: A subset $S \subseteq \{a_1, \ldots, a_n\}$ such that $\gcd(S) = \gcd(a_1, \ldots, a_n)$
MEASURE: The size $|S|$ of the subset S

Minimum GCD Multiplier in ℓ_∞-norm (MinGCDM$_\infty$)
INSTANCE: n numbers $a_1, \ldots, a_n \in \mathbb{Z}$
SOLUTION: A vector $\mathbf{x} \in \mathbb{Z}^n$ such that $\sum_{i=1}^{n} x_i a_i = \gcd(a_1, \ldots, a_n)$
MEASURE: The ℓ_∞-norm $\|\mathbf{x}\|_\infty := \max_{1 \le i \le n} |x_i|$ of the vector $\mathbf{x}$

Both problems have been shown to be **NP**-complete by Majewski and Havas [16].

For the MinGCDS problem we present a polynomial-time approximation algorithm achieving an approximation factor $1 + \ln n$ and prove that this algorithm is best possible. Specifically, we prove that unless $\mathbf{NP} \subseteq \mathbf{DTIME}(n^{O(\log \log n)})$ the optimum solution of MinGCDS cannot be approximated in polynomial-time within a factor $(1 - o(1)) \ln n$. Roughly speaking, the proof of both factors results from the similiarity between the problems MinGCDS and Min SetCover, where the latter **NP**-complete problem (see Karp [13] and Johnson [11]) admits the same approximability and inapproximability factor and is stated as follows:

Minimum Set Cover (MinSC)
INSTANCE: Finite set U and a collection of subsets $S_1, \ldots, S_m \subseteq U$ satisfying $\cup_{i=1}^{m} S_i = U$
SOLUTION: A subcollection $I \subseteq \{1, \ldots, m\}$ satisfying $\cup_{i \in I} S_i = U$
MEASURE: The number $|I|$ of sets in the subcollection

For the MinGCDM$_\infty$ problem, we are only able to derive less tight results on its approximability. From Majewski and Havas [16] we know that there exists a polynomial-time algorithm which computes a gcd multiplier $\mathbf{x} \in \mathbb{Z}^n$ for $a_1, \ldots, a_n$ with $\|\mathbf{x}\|_\infty \leq 0.5 \max_{1 \leq i \leq n} |a_i|$. We propose an algorithm which computes for a vector $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{Z}^n$ in polynomial-time a vector $\mathbf{x} \in \mathbb{Z}^n$ satisfying $\langle \mathbf{x}, \mathbf{a} \rangle = \gcd(a_1, \ldots, a_n)$ and $\|\mathbf{x}\| \leq 1.5^n \|\mathbf{a}\| / \gcd(a_1, \ldots, a_n)$.

On the other hand, we prove that, unless $\mathbf{NP} \subseteq \mathbf{DTIME}(n^{\mathrm{poly}(\log n)})$, MinGCDM$_\infty$ cannot be approximated within a factor $2^{\log^{1-\gamma} n}$, where γ is an arbitrary small positive constant. For the proof of this inapproximability result we construct a gap-preserving reduction from the Min Label Cover problem in ℓ_∞-norm to the MinGCDM$_\infty$ problem via the problems Min $\mathbb{Z}$-Solution of Linear System in ℓ_∞-norm and Minimum Diophantine Equation Solution in ℓ_∞-norm. The latter two minimization problems are defined as follows:

Min $\mathbb{Z}$-Solution of Linear System in ℓ_∞-norm (MinLS$_\infty$)
INSTANCE: A linear system $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ of m equations in n variables where $\boldsymbol{A}$ is a rational $m \times n$ matrix and $\mathbf{b}$ an n-dimensional rational vector
SOLUTION: A nonzero vector $\mathbf{x} \in \mathbb{Z}^n$ satisfying $\boldsymbol{A}\mathbf{x} = \mathbf{b}$
MEASURE: The ℓ_∞-norm $\|\mathbf{x}\|_\infty := \max_{1 \leq i \leq n} |x_i|$ of the vector $\mathbf{x}$

Minimum Diophantine Equation Solution in ℓ_∞-norm (MinDES$_\infty$)
INSTANCE: An equation $x_1 a_1 + \cdots + x_n a_n = b$ with $a_1, \ldots, a_n, b \in \mathbb{Z}$
SOLUTION: A vector $\mathbf{x} \in \mathbb{Z}^n$ such that $\sum_{i=1}^n x_i a_i = b$
MEASURE: The ℓ_∞-norm $\|\mathbf{x}\|_\infty := \max_{1 \leq i \leq n} |x_i|$ of the vector $\mathbf{x}$

The Min Label Cover problem in ℓ_∞-norm is defined in Section 4 and shown to be not approximable within a factor $2^{\log^{1-\gamma} n}$, where γ is an arbitrary small positive constant, unless $\mathbf{NP} \subseteq \mathbf{DTIME}(n^{\mathrm{poly}(\log n)})$. From the gap-preserving reductions we see that the same inapproximability factor holds for the problems MinLS$_\infty$ and MinDES$_\infty$.

2 Preliminaries

We briefly introduce some notation (see [6]).

Definition 1. An *optimization problem Π* is a set $\mathcal{I} \subseteq \{0,1\}^*$ of instances, a set $\mathcal{S} \subseteq \{0,1\}^*$ of feasible solutions on input $I \in \mathcal{I}$ and a polynomial-time computable measure $m : \mathcal{I} \times \mathcal{S} \to \mathbb{R}_+$, that assigns each tuple of instance I and solution S, a positive real number $m(I, S)$, called the *value* of the solution S. The optimization problem is to find, for a given input $I \in \mathcal{I}$ a solution $S \in \mathcal{S}$ such that $m(I, S)$ is optimum over all possible $S \in \mathcal{S}$.

If the optimum is $\min_{S \in \mathcal{S}}\{m(I,S)\}$ (resp. $\max_{S \in \mathcal{S}}\{m(I,S)\}$) we refer to Π as a *minimization* (resp. *maximization*) problem.

Definition 2. For an input I of a minimization (resp. maximization) problem Π whose optimal solution has value $opt_\Pi(I)$, an algorithm A is said to *approximate $opt_\Pi(I)$ within a factor $f(I)$* iff

$$opt_\Pi(I) \leq A(I) \leq f(I)\, opt_\Pi(I) \qquad (\text{resp. } opt_\Pi(I)/f(I) \leq A(I) \leq opt_\Pi(I)),$$

where $f(I) \geq 1$ and $A(I) > 0$.

For studying the hardness of approximation problems we introduce the following reduction due to Arora [2].

Definition 3. Let Π and Π' be two minimization problems and ρ, $\rho' \geq 1$. A *gap-preserving reduction from Π to Π' with parameters* $((c, \rho), (c', \rho'))$ is a polynomial-time transformation τ mapping every instance I of Π to an instance $I' = \tau(I)$ of Π' such that for the optima $opt_{\Pi}(I)$ and $opt_{\Pi'}(I')$ of I and I', respectively, the following holds:

$$opt_{\Pi}(I) \leq c \Longrightarrow opt_{\Pi'}(I') \leq c'$$
$$opt_{\Pi}(I) > c \cdot \rho \Longrightarrow opt_{\Pi'}(I') > c' \cdot \rho',$$

where c, ρ and c', ρ' depend on the instance sizes $|I|$ and $|I'|$, respectively.

3 Polynomial Time Approximation Algorithms for MinGCDS and MinGCDM$_2$

Main Theorem 4. *Given a set U of n numbers $a_1, \ldots, a_n \in \mathbb{Z}$, there exists a polynomial-time algorithm which computes a subset $S \subseteq \{a_1, \ldots, a_n\}$ satisfying* $\gcd(S) = \gcd(U)$ *and* $|S| \leq (1 + \ln n)\, opt_{\text{MinGCDS}}(U)$.

Proof. The following algorithm working in a greedy fashion will do the work.

```
input $\{a_1, \ldots, a_n\} =: U$
$S := \emptyset$
repeat
  choose an $x \in U \setminus S$ which minimizes $\gcd(S \cup \{x\})$
  $S := S \cup \{x\}$
until $\gcd(S) = \gcd(U)$
```

Obviously, the above algorithm runs in polynomial-time and computes a subset $S \subseteq \{a_1, \ldots, a_n\} = U$ with $\gcd(S) = \gcd(U)$. Thus, it remains to show that the set S satisfies $|S| \leq (1 + \ln n)\, opt_{\text{MinGCDS}}(U)$ as claimed above.

Since $opt_{\text{MinGCDS}}(U) = opt_{\text{MinGCDS}}(\{a_1/\gcd(U), \ldots, a_n/\gcd(U)\})$, we may assume that $\gcd(U) = 1$. With the primes $p_1, \ldots, p_s$ contained in the factorization of the numbers $a_1, \ldots, a_n$ we first define the numbers $\delta_{ij} \in \mathbb{N}_0$, $1 \leq i \leq n$, $1 \leq j \leq s$, as

$$\prod_{j=1}^{s} p_j^{\delta_{ij}} = a_i$$

and the numbers $M_j \in \mathbb{N}$, $1 \leq j \leq s$, as $M_j = \max_{1 \leq i \leq n} \delta_{ij}$, i.e.,

$$\prod_{j=1}^{s} p_j^{M_j} = \operatorname{lcm}(a_1, \ldots, a_n).$$

Consider now the *multiset* MinSC instance given by the universe $\mathcal{U}$, consisting in M_j copies of each prime p_j and the subsets $\mathcal{S}_i$, consisting in $M_j - \delta_{ij}$ copies of each prime p_j with $1 \leq i \leq n$ and $1 \leq j \leq s$. Since $\gcd(U) = \prod_{j=1}^{s} p_j^0$, it is obvious that computing a minimum gcd set for U amounts to find a minimum set cover for $\mathcal{U}$ among the subsets $\mathcal{S}_1, \ldots, \mathcal{S}_n$. Considering our algorithm we see that it follows the straightforward heuristic

$$\operatorname*{maximize}_{x \in U \setminus S} |\{\delta_j = 0 \mid \textstyle\prod_{j=1}^{s} p_j^{\delta_j} = \gcd(S \cup \{x\})\}|.$$

If I denotes the current index set with $\cup_{\iota \in I} \mathcal{S}_\iota := S$, this heuristic amounts to the heuristic

$$\operatorname*{maximize}_{i \in \{1,\ldots,n\} \setminus I} |\{j \mid \cup_{\iota \in I \cup \{i\}} \mathcal{S}_\iota \text{ contains } M_j \text{ copies of } p_j\}|$$

used for the MinSC approximation algorithm from Johnson [11], adapted to multisets. Therefore, every single selection-step of our algorithm is equivalent to a single selection-step of the multiset MinSC approximation algorithm from Johnson [11]. Thus, we may apply Chvatal's [7] analysis of the MinSC approximation algorithm. Chvatal [7] has shown that also the multiset version of the MinSC approximation algorithm achieves an approximation factor of $(1 + \ln n)$, which proves the claim. □

The following Theorem was implicitly proven by Håstad, Just, Lagarias and Schnorr [10, Sec. 6].

Main Theorem 5. *Given a vector* $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{Z}^n$*, there exists an algorithm which is polynomial in the bitlength of the input and computes an integral vector* $\mathbf{x} \in \mathbb{Z}^n$ *satisfying* $\langle \mathbf{x}, \mathbf{a} \rangle = \gcd(a_1, \ldots, a_n)$ *and* $\|\mathbf{x}\| \leq 1.5^n \|\mathbf{a}\| / \gcd(a_1, \ldots, a_n)$*.*

Proof. The main purpose of the algorithm presented in [10] was, in fact, to compute $n-1$ linearly independent integer relations for an input vector $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{Z}^n$, i.e., linearly independent integral vectors $\mathbf{m}_1, \ldots, \mathbf{m}_{n-1} \in \mathbb{Z}^n$ satisfying $\langle \mathbf{m}_i, \mathbf{a} \rangle = 0$, $i = 1, \ldots, n-1$. The algorithm extends the vector $\mathbf{a} / \gcd(a_1, \ldots, a_n) \in \mathbb{Z}^n$ to a generating system $\{\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_n\} = \{\mathbf{a} / \gcd(a_1, \ldots, a_n), \mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of the integral lattice $\mathbb{Z}^n$, where $\mathbf{e}_i$ is the i^{th} unit-vector in $\mathbb{Z}^n$. While leaving the vector $\mathbf{b}_0$ fixed, the algorithm applies the L^3-lattice basis reduction algorithm to the vectors $\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_n$.

The algorithm terminates with a basis $\mathbf{b}_0, \mathbf{b}_2, \ldots, \mathbf{b}_n$ of the lattice $\mathbb{Z}^n$ and $\mathbf{b}_1 = q\mathbf{b}_0$ for some $q \in \mathbb{Z}$. By lattice basis reduction theory the dual lattice vectors $\mathbf{c}_2, \ldots, \mathbf{c}_n$ which are the rows of the inverse matrix $[\mathbf{c}_1, \ldots, \mathbf{c}_n]^\top := [\mathbf{b}_0, \mathbf{b}_2, \ldots, \mathbf{b}_n]^{-1}$ form a basis of the lattice $\mathcal{L}_{\mathbf{a}}$ consisting of the zero vector $\mathbf{0}$ and all integer relations for $\mathbf{a}$. The analysis of the algorithm shows that the vector $\mathbf{c}_1 =: \mathbf{x}$, which is the only dual basis vector not orthogonal to $\mathbf{a}$, satisfies $\langle \mathbf{x}, \mathbf{a} \rangle = \gcd(a_1, \ldots, a_n)$. Moreover, from a closer look to the proofs given in [10] we infer $\|\mathbf{x}\| \leq 1.5^n \|\mathbf{a}\| / \gcd(a_1, \ldots, a_n)$; furthermore, the algorithm performs $\operatorname{poly}(n, \lceil \log \|\mathbf{a}\|_\infty \rceil)$ bit operations. □

4 Hardness of Min $\mathbb{Z}$-Solution of Linear System

4.1 The Min Label Cover Problem

In the following $G = (V_1, V_2, E)$ denotes a bipartite graph, $\mathcal{B}$ a set of labels for the vertices in $V_1 \cup V_2$, and for every $e \in E$ there exists a partial function $\Pi_e : \mathcal{B} \to \mathcal{B}$ describing the admissible pairs of labels. Moreover, we assume that G is d-regular, i.e., $|E| = d|V_1| = d|V_2|$. This property of G is a result of the reduction in [2, 3] from 3-Sat to Min Label Cover (see also Arora and Lund [4]). We adapt the notation of [2, 3].

Definition 6. A *labelling* of $G = (V_1, V_2, E)$ is a pair $(\mathcal{P}_1, \mathcal{P}_2)$ of functions $\mathcal{P}_i : V_i \to 2^{\mathcal{B}}$, $i = 1, 2$, assigning each vertex in $V_1 \cup V_2$ a possibly empty set of labels.

Definition 7. Let $(\mathcal{P}_1, \mathcal{P}_2)$ a labelling of $G = (V_1, V_2, E)$ and $e = (v_1, v_2)$, $v_1 \in V_1$, $v_2 \in V_2$, an edge of G. We call $e = (v_1, v_2)$ *covered* iff $\mathcal{P}_1(v_1) \neq \emptyset$, $\mathcal{P}_2(v_2) \neq \emptyset$ and for all labels $b_2 \in \mathcal{P}_2(v_2)$ there exists a label $b_1 \in \mathcal{P}_1(v_1)$ such that $\Pi_e(b_1) = b_2$. A labelling $(\mathcal{P}_1, \mathcal{P}_2)$ of $G = (V_1, V_2, E)$ is called a *total-cover* of G iff every edge of G is covered by the labelling $(\mathcal{P}_1, \mathcal{P}_2)$.

Definition 8. The ℓ_∞*-cost* of a labelling $(\mathcal{P}_1, \mathcal{P}_2)$ for a graph $G = (V_1, V_2, E)$ is defined as

$$cost(\mathcal{P}_1, \mathcal{P}_2) = \max_{v_1 \in V_1} |\mathcal{P}_1(v_1)|.$$

Definition 9. Min Label Cover in ℓ_∞-norm (MinLC$_\infty$)
INSTANCE: A d-regular bipartite graph $G = (V_1, V_2, E)$, a set of labels $\mathcal{B} = \{1, \ldots, \mathcal{N}\}$, $\mathcal{N} \in \mathbb{N}_+$, and for every edge $e \in E$ a partial function $\Pi_e : \mathcal{B} \to \mathcal{B}$ such that $\Pi_e^{-1}(1) \neq \emptyset$ for the distinguished label $1 \in \mathcal{B}$
SOLUTION: A total-cover $(\mathcal{P}_1, \mathcal{P}_2)$ of G
MEASURE: The ℓ_∞-cost $cost(\mathcal{P}_1, \mathcal{P}_2)$ of the total-cover $(\mathcal{P}_1, \mathcal{P}_2)$

Remark. In the above definition we can always ensure the existence of a total-cover with ℓ_∞-cost at most $\mathcal{N}$; we simply let $\mathcal{P}_2(v_2) = \{1\}$ for all $v_2 \in V_2$ and $\mathcal{P}_1(v_1) = \mathcal{B}$ for all $v_1 \in V_1$.

A slightly weaker form of the following Lemma is implicitly proved in Arora, Babai, Stern and Sweedyk [3].

Lemma 10. *There exists an almost-polynomial-time, i.e.,* $\mathbf{DTIME}(n^{\mathrm{poly}(\log n)})$, *transformation* τ *from* 3-Sat *to* Min Label Cover *such that, for all instances* I*:*

$$I \in \text{3-Sat} \Longrightarrow \exists \text{ total-cover } (\mathcal{P}_1, \mathcal{P}_2) \text{ of } \tau(I) : cost(\mathcal{P}_1, \mathcal{P}_2) = 1$$
$$I \notin \text{3-Sat} \Longrightarrow \forall \text{ total-cover } (\mathcal{P}_1, \mathcal{P}_2) \text{ of } \tau(I) : cost(\mathcal{P}_1, \mathcal{P}_2) > 2^{\log^{1-\gamma} |\tau(I)|},$$

where γ *is an arbitrary small positive constant.*

Proof. We have to make a detour around the maximization version of the label cover problem, which comes as follows:

MAX LABEL COVER (MAXLC)
INSTANCE: A d-regular bipartite graph $G = (V_1, V_2, E)$, a set of labels $\mathcal{B} = \{1, \ldots, \mathcal{N}\}$, $\mathcal{N} \in \mathbb{N}_+$, and for every edge $e \in E$ a partial function $\Pi_e : \mathcal{B} \to \mathcal{B}$ such that $\Pi_e^{-1}(1) \neq \emptyset$ for the distinguished label $1 \in \mathcal{B}$
SOLUTION: A labelling $(\mathcal{P}_1, \mathcal{P}_2)$ of G with $|\mathcal{P}_i(v_i)| \leq 1$ for all $v_i \in V_i$, $i = 1, 2$
MEASURE: The fraction of covered edges by $(\mathcal{P}_1, \mathcal{P}_2)$

Combining ideas of Arora, Babai, Stern and Sweedyk [3], Lund and Yannakakis [14] and the recent result of Raz [17], Arora and Lund [4] have shown, that there exists an almost-polynomial-time, i.e., $\mathbf{DTIME}(n^{\text{poly}(\log n)})$, transformation σ : 3-SAT → MAXLC such that, for all instances I:

$$I \in \text{3-SAT} \Longrightarrow opt_{\text{MaxLC}}(\sigma(I)) = 1$$
$$I \notin \text{3-SAT} \Longrightarrow opt_{\text{MaxLC}}(\sigma(I)) < 2^{-\log^{1-\gamma}|\sigma(I)|},$$

where γ is an arbitrary small positive constant. We will show that for any instance I of the label cover problem, we have

$$opt_{\text{MaxLC}}(I) \geq \frac{1}{opt_{\text{MinLC}_\infty}(I)}, \tag{1}$$

which proves the Theorem.

Consider a solution $(\mathcal{P}_1, \mathcal{P}_2)$ of a MINLC$_\infty$ instance I, placing at most $opt_{\text{MinLC}_\infty}(I)$ labels on any vertex in V_1. Randomly deleting all but one label for each vertex in $V_1 \cup V_2$ gives us a candidate solution $(\mathcal{P}_1', \mathcal{P}_2')$ for the corresponding MAXLC problem given by I. $(\mathcal{P}_1, \mathcal{P}_2)$ covers every edge $e = (u, v)$ of G and assigns for every label b_2 assigned to $v \in V_2$ a label a_1 to $u \in V_1$ such that $\Pi_e(b_1) = b_2$. Thus, the expected number of edges in the randomly constructed labelling $(\mathcal{P}_1', \mathcal{P}_2')$ is at least

$$\sum_{e=(u,v)} |\mathcal{P}_1(u)|^{-1} \geq \frac{|E|}{opt_{\text{MinLC}_\infty}(I)}.$$

Therefore, the expected fraction of covered edges by $(\mathcal{P}_1', \mathcal{P}_2')$ is at least $1/opt_{\text{MinLC}_\infty}(I)$. Hence, there must exists a solution covering at least $|E|/opt_{\text{MinLC}_\infty}(I)$ edges, showing equation (1).

□

4.2 MIN $\mathbb{Z}$-SOLUTION OF LINEAR SYSTEM

The following result transforms the inapproximability gap from MIN LABEL COVER to MIN $\mathbb{Z}$-SOLUTION OF LINEAR SYSTEM.

Theorem 11. *There exists a polynomial-time transformation τ from* Min Label Cover *to* Min $\mathbb{Z}$-Solution of Linear System *such that, for all instances I and for all $\rho \geq 1$:*

$$opt_{\mathrm{MinLC}_\infty}(I) = 1 \Longrightarrow opt_{\mathrm{MinLS}_\infty}(\tau(I)) = 1$$
$$opt_{\mathrm{MinLC}_\infty}(I) > \rho \Longrightarrow opt_{\mathrm{MinLS}_\infty}(\tau(I)) > \sqrt{\rho}.$$

Proof. From a given Min Label Cover instance $I = (V_1, V_2, E, \Pi, \mathcal{B}, \mathcal{N})$ we construct a linear system of equations $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ with $\boldsymbol{A}$ an $m \times n$ matrix of entries $\{-1, 0, 1\}$, $\mathbf{b}$ an m-dimensional vector of entries $\{0, 1\}$, $m = |V_1|\mathcal{N} + |E|(\mathcal{N} + 1)$ and $n = 2|V_1|\mathcal{N} + |V_2|\mathcal{N}$.

For every pair (v, b) with $v \in V_1 \cup V_2$ and $b \in \mathcal{B}$ we define a column vector $\mathbf{a}_{v,b} \in \{-1, 0, 1\}^m$ of $\boldsymbol{A}$ as follows. The first $|E|(\mathcal{N} + 1)$ coordinates of $\mathbf{a}_{v,b}$ are split into $|E|$ blocks of *e-projections* $\mathbf{u}_e(\mathbf{a}_{v,b})$ — one $(\mathcal{N} + 1)$-length block for every edge $e \in E$. In particular, we define for every $(v_2, b_2) \in V_2 \times \mathcal{B}$

$$\mathbf{u}_e(\mathbf{a}_{v_2,b_2}) := \begin{cases} \mathbf{e}_{b_2} & \text{iff } e \text{ is incident to } v_2 \\ \mathbf{0} & \text{otherwise} \end{cases}$$

and for every $(v_1, b_1) \in V_1 \times \mathcal{B}$

$$\mathbf{u}_e(\mathbf{a}_{v_1,b_1}) := \begin{cases} \mathbf{1} - \mathbf{e}_{\Pi_e(b_1)} & \text{iff } e \text{ is incident to } v_1 \text{ and } \Pi_e(b_1) \neq \emptyset \\ \mathbf{0} & \text{otherwise} \end{cases}$$

where $\mathbf{e}_j$, $j = 1, \ldots, \mathcal{N}$, denotes the j^{th}-unit vector and $\mathbf{0}$, $\mathbf{1}$ the all-zero, all-one vector in $\mathbb{R}^{\mathcal{N}+1}$, respectively.

The definition of the remaining $|V_1|\mathcal{N}$ coordinates of $\mathbf{a}_{v,b}$ uses the properties of Hadamard matrices. A *Hadamard matrix of order ℓ*, denoted by $\boldsymbol{H}_\ell$, is an $\ell \times \ell$ matrix with ± 1 entries such that $\boldsymbol{H}_\ell \boldsymbol{H}_\ell^\top = \ell \boldsymbol{I}_\ell$. The columns of $\frac{1}{\sqrt{\ell}}\boldsymbol{H}_\ell$ clearly form an orthonormal basis. Therefore $\|\frac{1}{\sqrt{\ell}}\boldsymbol{H}_\ell \mathbf{z}\|_2 = \|\mathbf{z}\|_2$ for every $\mathbf{z} \in \mathbb{Z}^\ell$. If $\mathbf{z} \in \mathbb{Z}^\ell$ has at least k nonzero entries we thus have $\|\boldsymbol{H}_\ell \mathbf{z}\|_\infty \geq \sqrt{k}$. Hadamard matrices can be constructed in time linear in the size of the matrix if ℓ is a power of 2, cf. [15]. Otherwise we use the matrix $\boldsymbol{H}_\ell$ consisting in the first ℓ columns of the Hadamard matrix of order $2^{\lceil \log \ell \rceil}$. Clearly, $\|\boldsymbol{H}_\ell \mathbf{z}\|_\infty \geq \sqrt{k}$ remains valid for vectors $z \in \mathbb{Z}^\ell$ with at least k nonzero entries.

We may assume that for $\ell = \mathcal{N}$ there exists a Hadamard matrix $\boldsymbol{H}_\ell = [\mathbf{h}_1, \ldots, \mathbf{h}_\ell]$ with column vectors $\mathbf{h}_b$ of $\boldsymbol{H}_\ell$, each of them uniquely identified with a label $b \in \mathcal{B}$. We now split the last $|V_1|\mathcal{N}$ coordinates of $\mathbf{a}_{v,b}$ into $|V_1|$ blocks of *v_1-projections* $\mathbf{u}_{v_1}(\mathbf{a}_{v,b})$ — one $\mathcal{N}$-length block for every vertex $v_1 \in V_1$ — where the v_1-projections for every $v \in V_1 \cup V_2$ and $b \in \mathcal{B}$ are defined as follows

$$\mathbf{u}_{v_1}(\mathbf{a}_{v,b}) := \begin{cases} \mathbf{h}_b & \text{iff } v = v_1 \\ \mathbf{0} & \text{otherwise} \end{cases}$$

and $\mathbf{0}$ denotes the all-zero vector in $\mathbb{R}^{\mathcal{N}}$. This definition clearly implies $\mathbf{u}_{v_1}(\mathbf{a}_{v,b}) = \mathbf{0}$ for all $v \in V_2$ and all $b \in \mathcal{B}$.

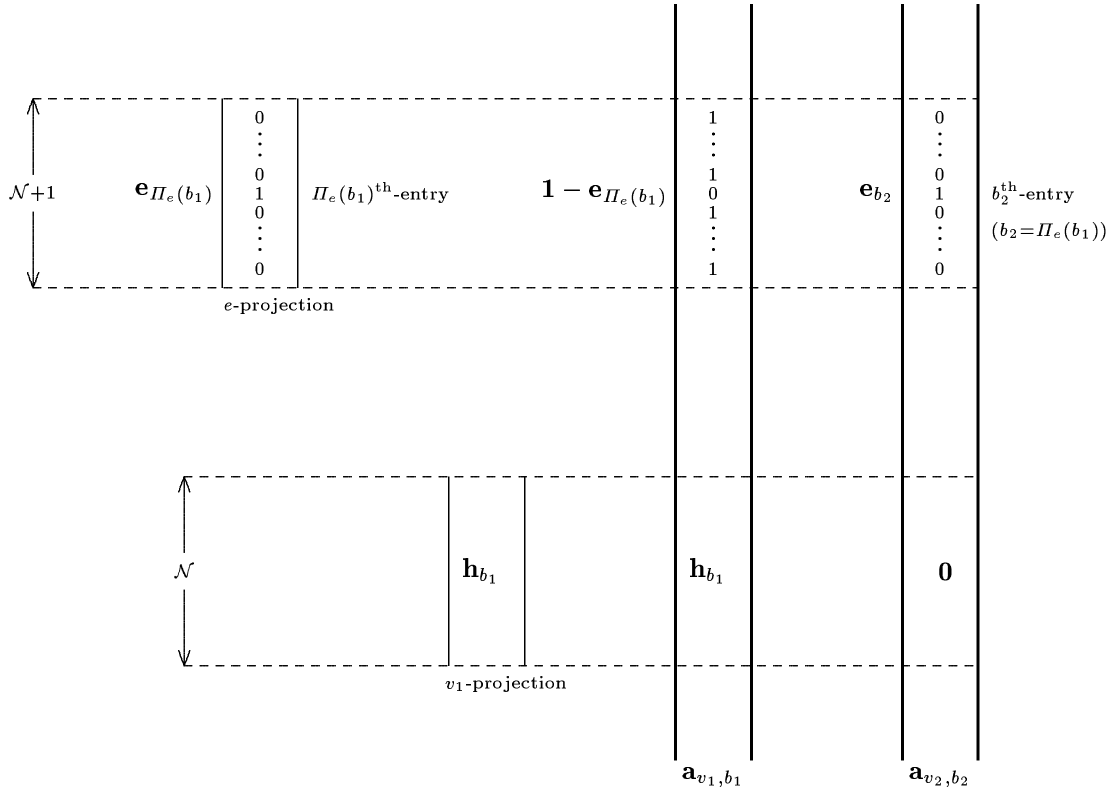

Fig. 1. The resulting column vectors due to [ABS$^+$93]

Moreover, the remaining $|V_1|\mathcal{N}$ column vectors are $\mathbf{e}_{|E|(\mathcal{N}+1)+i}$, $i = 1,\dots,|V_1|\mathcal{N}$, where $\mathbf{e}_j$ denotes the j^{th}-unit vector in $\mathbb{R}^{|V_1|\mathcal{N}+|E|(\mathcal{N}+1)}$.

Eventually, we define the right hand side of our linear system — the vector $\mathbf{b}$ — as the vector having 1 in each of the first $|E|(\mathcal{N}+1)$ coordinates and 0 in the remaining ones.

$$\begin{array}{rccc} & |V_1|\mathcal{N}\text{ columns} & |V_2|\mathcal{N}\text{ columns} & |V_1|\mathcal{N}\text{ columns} \\ |E|(\mathcal{N}+1)\text{ rows} & [\mathbf{u}_e(\mathbf{a}_{v,b_1}),\dots,\mathbf{u}_e(\mathbf{a}_{v,b_{\mathcal{N}}})]_{\substack{e\in E\\ v\in V_1}} & [\mathbf{u}_e(\mathbf{a}_{v,b})]_{\substack{e\in E\\ (v,b)\in V_2\times\mathcal{B}}} & \mathbf{0} \\ |V_1|\mathcal{N}\text{ rows} & [\mathbf{u}_{v_1}(\mathbf{a}_{v,b_1}),\dots,\mathbf{u}_{v_1}(\mathbf{a}_{v,b_{\mathcal{N}}})]_{\substack{v_1\in V_1\\ v\in V_1}} & \mathbf{0} & \boldsymbol{I}_{|V_1|\mathcal{N}} \end{array} \mathbf{x} = \begin{pmatrix}\mathbf{1}\\ \\ \mathbf{0}\end{pmatrix}$$

Fig. 2. The linear system $\boldsymbol{A}\mathbf{x} = \mathbf{b}$

$\Big($owing lack of space we abbreviated $[\mathbf{u}_e(\mathbf{a}_{v,b_1}),\dots,\mathbf{u}_e(\mathbf{a}_{v,b_{\mathcal{N}}})]_{\substack{e\in E\\ v\in V_2}} =: [\mathbf{u}_e(\mathbf{a}_{v,b})]_{\substack{e\in E\\ (v,b)\in V_2\times\mathcal{B}}}\Big)$

Given a vector $\mathbf{y} \in \mathbb{R}^{|V_1|\mathcal{N}+|E|(\mathcal{N}+1)}$ let $\mathbf{u}_E(\mathbf{y})$ denote the vector $\mathbf{y}$ *restricted to* its first $|E|(\mathcal{N}+1)$ coordinates. Let $\mathbf{x} = \sum x_{v,b}\mathbf{u}_E(\mathbf{a}_{v,b})$ be an integral linear combination of the 'restricted' column vectors $\mathbf{u}_E(\mathbf{a}_{v,b})$. Then, assigning every vertex v a label b iff $x_{v,b} \neq 0$ defines a labelling $(\mathcal{P}_1^{\mathbf{x}}, \mathcal{P}_2^{\mathbf{x}})$ *induced by the vector* $\mathbf{x}$. From [3, Corollary 10] it follows that any such $\mathbf{x}$ with $\mathbf{x} = \mathbf{u}_E(\mathbf{b})$, induces a total-cover of (V_1, V_2, E).

Thus, any solution $\mathbf{x} \in \mathbb{Z}^{2|V_1|\mathcal{N}+|V_2|\mathcal{N}}$ of the linear system $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ induces by its first $|V_1|\mathcal{N} + |V_2|\mathcal{N}$ coordinates a total-cover of (V_1, V_2, E) (note that the last $|V_1|\mathcal{N}$ column vectors of the matrix $\boldsymbol{A}$ have 0-entries in its first $|E|(\mathcal{N}+1)$ coordinates).

Thus, for the induced total-cover $(\mathcal{P}_1^{\mathbf{x}}, \mathcal{P}_2^{\mathbf{x}})$ there exists a vertex, say $v_1 \in V_1$, with at least $opt_{\mathrm{MinLC}_\infty}(I)$ labels assigned. This in turn means that $\mathbf{x}$ has at least $opt_{\mathrm{MinLC}_\infty}(I)$ nonzero entries. By the above properties of the Hadamard matrices we see that there exists an index $i^* \in \{|E|(\mathcal{N}+1)+1, \ldots, |E|(\mathcal{N}+1) + |V_1|\mathcal{N}\}$ such that

$$\left| \sum_{j=1}^{|V_1|\mathcal{N}} a_{i^*,j} x_j \right| \geq \sqrt{opt_{\mathrm{MinLC}_\infty}(I)}.$$

As $\mathbf{x}$ is a solution of $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ its remaining $|V_1|\mathcal{N}$ coordinates are forced to cancel out each of the sums

$$\sum_{j=1}^{|V_1|\mathcal{N}} a_{i,j} x_j, \qquad i = |E|(\mathcal{N}+1)+1, \ldots, |E|(\mathcal{N}+1) + |V_1|\mathcal{N}.$$

Hence, any solution $\mathbf{x}$ of $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ has one entry, say $x_{|V_1|\mathcal{N}+|V_2|\mathcal{N}+j^*}$, $j^* \in \{1, \ldots, |V_1|\mathcal{N}\}$ satisfying

$$\|\mathbf{x}\|_\infty \geq |x_{|V_1|\mathcal{N}+|V_2|\mathcal{N}+j^*}| \geq \sqrt{opt_{\mathrm{MinLC}_\infty}(I)}.$$

Now assume $opt_{\mathrm{MinLC}_\infty}(I) = 1$. Let $(\mathcal{P}_1, \mathcal{P}_2)$ denote the corresponding labelling. Then, the $(2|V_1|\mathcal{N} + |V_2|\mathcal{N})$-length vector $\mathbf{x}$ given by

$$\begin{aligned} x_{v_i,\mathcal{P}_i(v_i)} &:= 1 \quad && \forall v_i \in V_i,\ i = 1,2 \\ x_{v_i,b} &:= 0 && \forall v_i \in V_i,\ \forall b \in \mathcal{B} \setminus \mathcal{P}_i(v_i),\ i = 1,2 \\ x_{|V_1|\mathcal{N}+|V_2|\mathcal{N}+1+i} &:= -x_i && i = 1, \ldots, |V_1|\mathcal{N} \end{aligned}$$

obviously is a feasible solution of the linear system $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ satisfying $\|\mathbf{x}\|_\infty = 1$.

The reduction from the given instance I of Min Label Cover to the above constructed linear system $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ is feasible in time polynomial in the dimension of $\boldsymbol{A}$ which in turn is polynomial in $|I|$. Clearly, the above reduction, say τ, is gap-reserving with parameters $((1, \rho), (1, \sqrt{\rho}))$. □

The above Theorem implies the following.

Corollary 12. *Approximating* MinLS$_\infty$ *within a factor* $2^{\log^{1-\gamma} n}$ *is almost-***NP***-hard, where* γ *is an arbitrary small positive constant and* n *is the size of the* MinLS$_\infty$*-instance.*

5 Hardness of Approximating MinGCDM$_\infty$

5.1 Aggregation — Part I

The following Lemma implicitely proven by Kannan [12] establishes a polynomial-time reduction from a system of inhomogeneous linear equations to a single equation with identical solution set, provided that the solutions are bounded.

Lemma 13. *Let $\boldsymbol{A}$ be an integral $m \times n$ matrix, $\|\boldsymbol{A}\|_\infty$ the maximum absolute value of its entries a_{ij}, $1 \leq i \leq m$, $1 \leq j \leq n$ and $\mathbf{b}$ be an integral m-dimensional vector. Then*

$$B_\mu \cap \{\mathbf{x} \in \mathbb{Z}^n \mid \boldsymbol{A}\mathbf{x} = \mathbf{b}\} = B_\mu \cap \left\{\mathbf{x} \in \mathbb{Z}^n \;\middle|\; \sum_{i=1}^{m}\sum_{j=1}^{n} k^i a_{ij} x_j = \sum_{i=1}^{m} k^i b_i\right\}$$

where B_μ denotes the n-dimensional ball of ℓ_∞-radius μ centered at the origin and $k = n\|\boldsymbol{A}\|_\infty \mu + \|\mathbf{b}\|_\infty + 1$.

Proof. Denote the two sets by S_m and S_1, respectively. Clearly, $S_m \subseteq S_1$. For proving the reverse inclusion suppose, that there exists an element $\mathbf{x} \in S_1$ not satisfying at least one equation of $\boldsymbol{A}\mathbf{x} = \mathbf{b}$. Let $\boldsymbol{A} =: [\mathbf{a}_1, \ldots, \mathbf{a}_n]^\top$ and let $i_{\max}$ denote the largest index for which $\langle \mathbf{a}_i, \mathbf{x}\rangle \neq b_i$. As $\|\mathbf{x}\|_\infty \leq \mu$ we have

$$|\langle \mathbf{a}_i, \mathbf{x}\rangle - b_i| \leq n\|\boldsymbol{A}\|_\infty \mu + \|\mathbf{b}\|_\infty = k - 1$$

and since $\mathbf{x} \in S_1$ we must have

$$\sum_{i=1}^{m} k^i(\langle \mathbf{a}_i, \mathbf{x}\rangle - b_i) = 0.$$

By definition of $i_{\max}$ this yields

$$\sum_{i=1}^{i_{\max}-1} k^i(\langle \mathbf{a}_i, \mathbf{x}\rangle - b_i) = -k^{i_{\max}}(\langle \mathbf{a}_{i_{\max}}, \mathbf{x}\rangle - b_{i_{\max}})$$

with a nonzero right-hand side implying that the left-hand side is also nonzero. Now the left-hand side is both a multiple of $k^{i_{\max}}$ and in absolute value bounded by $k^{i_{\max}} - k \leq k^{i_{\max}} - 1$, a contradiction. □

5.2 Hardness of Approximating Minimum Diophantine Equation Solution

Theorem 14. *There exists a polynomial-time transformation τ from* Min $\mathbb{Z}$-Solution of Linear System *to* Minimum Diophantine Equation Solution *such that the following holds:*

1. *for all instances I and for all $\rho \geq 1$:*

$$opt_{\mathrm{MinLS}_\infty}(I) = 1 \Longrightarrow opt_{\mathrm{MinDES}_\infty}(\tau(I)) = 1$$
$$opt_{\mathrm{MinLS}_\infty}(I) > \rho \Longrightarrow opt_{\mathrm{MinDES}_\infty}(\tau(I)) > \rho$$

2. the constructed instance $a'_1 x_1 + \cdots + a'_n x_n = b'$ *of* MinDES$_\infty$ *satisfies*

$$x a'_{j^*} = b' \textit{ for some } j^* \in \{1,\ldots,n\} \textit{ and some } x \in \mathbb{Z}.$$

Proof. Consider the linear system $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ constructed in the reduction of the proof of Theorem 11, see Figure 2. Recall that for the underlying d-regular graph $G = (V_1, V_2, E)$ with label set $\mathcal{B}$, $|\mathcal{B}| = \mathcal{N}$, the matrix $\boldsymbol{A}$ is a $m \times n$ matrix of entries $\{-1, 0, 1\}$ with $m = |V_1|\mathcal{N} + |E|(\mathcal{N}+1)$ and $n = 2|V_1|\mathcal{N} + |V_2|\mathcal{N}$, and that the vector $\mathbf{b}$ has only 1-entries in its first $|E|(\mathcal{N}+1)$ coordinates and only 0-entries in the remaining $|V_1|\mathcal{N}$ coordinates.

Thus, permuting the first $|E|(\mathcal{N}+1)$ rows of $\boldsymbol{A}$ does not change the solution set of the linear system $\boldsymbol{A}\mathbf{x} = \mathbf{b}$. Let $\mathbf{a}_{j^*} := \mathbf{a}_{v_2,b_2}$ for some $v_2 \in V_2$, $b_2 \in \mathcal{B}$ be a non-zero column vector of $\boldsymbol{A}$. Since the graph G is d-regular, $\mathbf{a}_{j^*}$ has exactly d 1-entries. Let σ be the permutation shifting all 1-entries of $\mathbf{a}_{j^*}$ in its first d coordinates and note that $\sigma(\mathbf{b}) = \mathbf{b}$.

Aggregating now the (solution equivalent) linear system $\sigma(\boldsymbol{A})\mathbf{x} = \mathbf{b}$ via Lemma 13 yields the inhomogenous Diophantine equation

$$\sum_{j=1}^{2|V_1|\mathcal{N}+|V_2|\mathcal{N}} a'_j x_j = b'$$

with

$$a'_j := \sum_{i=1}^{|V_1|\mathcal{N}+|E|(\mathcal{N}+1)} k^i a_{\sigma(i),j} \qquad \text{and} \qquad b' := \sum_{i=1}^{|V_1|\mathcal{N}+|E|(\mathcal{N}+1)} k^i b_i.$$

Particularly, we have

$$a'_{j^*} = \sum_{i=1}^{d} k^i = k\frac{k^d - 1}{k-1} \qquad \text{and} \qquad b' = \sum_{i=1}^{|E|(\mathcal{N}+1)} k^i = k\frac{k^{|E|(\mathcal{N}+1)} - 1}{k-1},$$

and by the regularity of G, i.e., $d|V_2| = |E|$, it follows that

$$k^d - 1 \mid k^{|E|(\mathcal{N}+1)} - 1, \qquad \text{hence} \qquad a'_{j^*} \mid b'.$$

Thus, we have constructed a gap-preserving reduction from the linear system $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ to an inhomogenous Diophantine equation instance $a'_1 x_1 + \cdots + a'_n x_n = b'$ with parameters $((1,\rho),(1,\rho))$ such that there is an index $j^* \in \{1,\ldots,n\}$ satisfying $a'_{j^*} \mid b'$. □

Corollary 15. *Approximating* MinDES$_\infty$ *within a factor* $2^{\log^{1-\gamma} n}$ *is almost-***NP***-hard, where* γ *is an arbitrary small positive constant and* n *is the size of the* MinDES$_\infty$*-instance.*

5.3 Aggregation — Part II

The following Lemma, originally proved by Anthonisse [1], is a slight variation of the former Lemma and crucial for our reduction; for the simple proof see also [18, 16].

Lemma 16. *Let $\boldsymbol{A}$ be an integral $2 \times n$ matrix and $b \in \mathbb{Z}$. Then*

$$B_\mu \cap \left\{ \mathbf{x} \in \mathbb{Z}^n \;\middle|\; \boldsymbol{A}\mathbf{x} = \left[\begin{smallmatrix} b \\ 0 \end{smallmatrix}\right] \right\} = B_\mu \cap \left\{ \mathbf{x} \in \mathbb{Z}^n \;\middle|\; \sum_{j=1}^{n} (a_{1,j} + k a_{2,j}) x_j = b \right\}$$

where B_μ denotes the n-dimensional ball of ℓ_∞-radius μ centered at the origin and $k > \mu \sum_{j=1}^{n} |a_{1,j}| + b$.

5.4 The Final Reduction

By piecing the above results together we now prove the following:

Main Theorem 17. *Unless* $\mathbf{NP} \subseteq \mathbf{DTIME}(n^{\mathrm{poly}(\log n)})$, *there exists no polynomial-time algorithm approximating the* Minimum GCD Multiplier *problem in ℓ_∞-norm within a factor of $2^{\log^{1-\gamma} n}$, where γ is an arbitrary small positive constant.*

Proof. We start with the instance $a_1' x_1 + \cdots + a_n' x_n = b'$ of MinDES$_\infty$ constructed in the above Theorem 14 and consider for an arbitrary integer $c \in \mathbb{Z} \setminus \{0\}$ the linear system

$$\begin{aligned} c x_{n+1} &= c \\ a_1' x_1 + \cdots + a_n' x_n - b' x_{n+1} &= 0 \end{aligned}$$

which enforces the variable x_{n+1} to take on the value 1. Now, we fix $\rho \geq 1$, choose $k = \rho c + c + 1$ and apply Lemma 16 to this linear system, obtaining the single equation

$$k a_1' x_1 + \cdots + k a_n' x_n + (c - k b') x_{n+1} = c.$$

We observe that the right hand side c in the last equation was an arbitrary chosen integer and that b' satisfies by Theorem 14

$$x a_{i^*}' = b' \text{ for some } i^* \in \{1, \ldots, n\} \text{ and some } x \in \mathbb{Z}\ . \qquad (*)$$

This will give us the desired gap-preserving reduction τ. Namely, we choose $c = \gcd(a_1', \cdots, a_n')$. By $(*)$ this implies

$$\begin{aligned} &\gcd(k a_1', \cdots, k a_n', (c - k b')) \\ &= \gcd(k a_1', \cdots, k a_{i^*-1}', k a_{i^*+1}', \ldots, k a_n', \gcd(k a_{i^*}', (c - k b'))) \\ &= \gcd(k a_1', \cdots, k a_{i^*-1}', k a_{i^*+1}', \ldots, k a_n', \gcd(k a_{i^*}', (c - k x a_{i^*}'))) \\ &= \gcd(k a_1', \cdots, k a_{i^*-1}', k a_{i^*+1}', \ldots, k a_n', \gcd(k a_{i^*}', c)) \\ &= \gcd(k \gcd(a_1', \cdots, a_{i^*-1}', a_{i^*}', a_{i^*+1}', \ldots, a_n'), c) \\ &= c. \end{aligned}$$

Since the variable x_{n+1} is enforced to take on the value 1, it is obvious from the above, Lemma 10, and Theorem 11 that the reduction τ from 3-SAT to MIN GCDM$_\infty$ with $\tau(I) = \{ka'_1, \cdots, ka'_n, (c - kb')\}$ satisfies for all instances I and all $\gamma > 0$

$$I \in \text{3-SAT} \Longrightarrow opt_{\mathrm{MinGCDM}_\infty}(\tau(I)) = 1$$
$$I \notin \text{3-SAT} \Longrightarrow opt_{\mathrm{MinGCDM}_\infty}(\tau(I)) > 2^{\log^{1-\gamma}|\tau(I)|}.$$

Therefore, given a polynomial-time algorithm approximating the MIN GCDM$_\infty$ problem within a factor of $2^{\log^{1-\gamma} n}$ for some $\gamma > 0$ would enable us to decide 3-SAT in almost-polynomial-time. □

6 Hardness of Approximating MINGCDS

The following very recent Lemma — indicating an approximation threshold of $\ln n$ for MINIMUM SET COVER — is due to Feige [8], which in turn is based on the work of Lund and Yannakakis [14].

Lemma 18. *There exists an almost-polynomial-time, i.e.,* **DTIME**$(n^{O(\log\log n)})$, *transformation* τ : 3-SAT → MINIMUM SET COVER *such that, for all instances* I*:*

$$I \in \text{3-SAT} \Longrightarrow opt_{\mathrm{MinSC}}(\tau(I)) = K(|I|)$$
$$I \notin \text{3-SAT} \Longrightarrow opt_{\mathrm{MinSC}}(\tau(I)) > (1 - o(1)) \cdot \ln(N) \cdot K(|I|),$$

where $K(|I|)$ *is a polynomial-time computable function and* N *the size of the ground set of the set cover instance* $\tau(I)$.

Theorem 19. *There exists a polynomial-time transformation* τ *from* MINIMUM SET COVER *to* MINIMUM GCD SET *such that, for all instances* I *and for all* $c, \rho \geq 1$*:*

$$opt_{\mathrm{MinSC}}(I) \leq c \Longrightarrow opt_{\mathrm{MinGCDS}}(\tau(I)) \leq c$$
$$opt_{\mathrm{MinSC}}(I) > c \cdot \rho \Longrightarrow opt_{\mathrm{MinGCDS}}(\tau(I)) > c \cdot \rho.$$

Proof. We give a gap-preserving reduction with parameters $((c, \rho), (c, \rho))$ from MINIMUM GCD SET to MINIMUM SET COVER which is due to Majewski and Havas [16]. Assume, we are given a MINIMUM SET COVER instance, i.e., a finite set U and a collection of subsets $S_1, \ldots, S_m \subseteq U$ with $S_i = \{s_{i_1}, \ldots, s_{i_d}\}$, $d > 0$, satisfying $\cup_{i=1}^m S_i = U$.

We select the first $n := |U|$ primes $p_1, \ldots, p_n$ and define with

$$\delta_{ij} := \begin{cases} 0 \text{ iff } s_j \in S_i, & i = 1, \ldots, m,\ j = 1, \ldots, n \\ 1 \text{ otherwise} \end{cases}$$

the numbers

$$a_i := \prod_{j=1}^{n} p_j^{\delta_{ij}}, \quad i = 1, \dots, m.$$

The greatest common divisor of $a_1, \dots, a_n$ is by definition

$$\prod_{j=1}^{n} p_j^{\min_{1 \le i \le m} \delta_{ij}} = \prod_{j=1}^{n} p_j^0 = 1.$$

Note that by the above '$s_j \notin S_i$' transforms into 'p_j is a prime factor of a_i' and vice versa. Therefore, every gcd set S with $\gcd(S) = 1$ yields a solution of size $|S|$ for a given MINIMUM SET COVER instance. This shows that the reduction is gap-preserving with parameters $((c, \rho), (c, \rho))$. Moreover, the reduction is polynomial in the size of the MINIMUM SET COVER instance as the bitlength of every number a_i is bounded by $n^{1+o(1)}$. □

By Lemma 18 and Theorem 19, we have established an almost-polynomial-time transformation from 3-SAT to MINIMUM GCD SET, which implies 3-SAT $\in$ $\mathbf{DTIME}(n^{O(\log \log n)})$, if there exists a polynomial-time algorithm approximating MINGCDS within a factor $(1 - o(1)) \ln n$. We thus conclude with the following:

Main Theorem 20. *Unless* $\mathbf{NP} \subseteq \mathbf{DTIME}(n^{O(\log \log n)})$, *there exists no polynomial-time algorithm approximating the* MINIMUM GCD SET *problem within a factor* $(1 - o(1)) \ln n$.

References

1. J. M. Anthonisse. A note on equivalent systems of linear diophantine equations. *Z. Operations Research*, Volume 17, pages 167–177, 1973.
2. S. Arora. *Probabilistic Checking of Proofs and Hardness of Approximation Problems.* Ph.D. thesis, University of California at Berkeley, 1994.
3. S. Arora, L. Babai, J. Stern and Z Sweedyk. The hardness of approximate optima in lattices, codes and systems of linear equations. In *Proc. 34th IEEE Symp. on Foundations of Computer Science*, pages 724–730, 1993.
4. S. Arora and C. Lund. Hardness of approximation. In D. Hochbaum (editor), *Approximation Algorithms for NP-hard Problems*, Chapter 11. PWS Publishing, 1996.
5. S. Arora, C. Lund, R. Motwani, M. Sudan and M. Szegedy. Proof verification and hardness of approximation problems. In *Proc. 33rd IEEE Symp. on Foundations of Computer Science*, pages 14–23, 1992.
6. G. Ausiello, P. Crescenzi and M. Protasi. Approximate solutions of NP optimization problems. *Theoretical Computer Science*, Volume 150, pages 1–55, 1995.
7. V. Chvatal. A greedy heuristic for the set-covering problem. *Mathematics of Operations Research*, Volume 4, pages 233–235, 1978.
8. U. Feige. A threshold of $\ln n$ for approximating set cover. In *Proc. 28th ACM Symp. Theory of Computing*, 1996.

9. U. Feige and L. Lovász. Two-prover one-round proof systems: Their power and their problems. In *Proc. 24th ACM Symp. Theory of Computing*, pages 643–654, 1992.
10. J. Håstad, B. Just, J. C. Lagarias and C. P. Schnorr. Polynomial time algorithms for finding integer relations among real numbers. *SIAM J. Computation*, Volume 18, pages 859–881, 1989.
11. D. S. Johnson. Approximation algorithms for combinatorial algorithms. *J. CSS*, Volume 9, pages 256–278, 1974.
12. R. Kannan. Polynomial-time aggregation of integer programming problems. *J. ACM*, Volume 30, pages 133–145, 1983.
13. R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher (editors), *Complexity of Computer Computations.* Plenum Press, New York, 1972.
14. C. Lund and M. Yannakakis. On the hardness of minimization problems. *J. ACM*, Volume 41, pages 960–981, 1994.
15. F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes.* North Holland, Amsterdam, 1977.
16. B. S. Majewski and G. Havas. The complexity of greatest common divisor computations. In *Proc. 1st International Symposium on Algorithmic Number Theory*, pages 184–193. Springer, 1994. LNCS 877.
17. R. Raz. A parallel repetition theorem. In *Proc. 27th ACM Symp. Theory of Computing*, pages 447–456, 1995.
18. P. van Emde Boas. Another NP-complete partition problem and the complexity of computing short vectors in a lattice. Technical Report 81-04, Math. Inst., University of Amsterdam, 1981.

This article was processed using the LaTeX macro package with LLNCS style