On Approximation Complexity of Metric Dimension Problem

Mathias Hauptmann* Richard Schmied† Claus Viehmann‡

Abstract

We study the approximation complexity of the *Metric Dimension problem* in bounded degree, dense as well as in general graphs. For the general case, we prove that the Metric Dimension problem is not approximable within $(1-\epsilon)\ln n$ for any $\epsilon > 0$, unless $NP \subseteq DTIME(n^{\log\log n})$, and we give an approximation algorithm which matches the lower bound.

Even for bounded degree instances it is APX-hard to determine (compute) the value of the metric dimension which we prove by constructing an approximation preserving reduction from the bounded degree Vertex Cover problem.

The special case, in which the underlying graph is superdense turns out to be APX-complete. In particular, we present a greedy constant factor approximation algorithm for these kind of instances and construct an approximation preserving reduction from the bounded degree Dominating Set problem. We also provide first explicit approximation lower bounds for the Metric Dimension problem restricted to dense and bounded degree graphs.

1 Introduction

In a connected graph $G = (V, E)$, a vertex $v \in V$ *resolves* or *distinguishes* a pair $u, w \in V$ if $d(v, u) \neq d(v, w)$, where $d(\cdot, \cdot)$ denotes the length of a shortest path between two vertices in G. A *resolving set* of G is a subset $V' \subseteq V$ such that for each pair $u, w \in V$ there exists some $v \in V'$ that distinguishes u and w. The minimum cardinality of a resolving set is called the *metric dimension* of G, denoted by $dim(G)$. The *Metric Dimension* problem asks to find a resolving set of minimum cardinality. We call here a graph $G = (V, E)$ k-superdense if the degree of every vertex is at least $|V| - k$ where k is a constant. Throughout the paper, we will use the notation $n := |V|$.

1.1 Related Work

The notion of resolving sets were introduced independently by Harary and Melter [13] and Slater [19]. Applications of resolving sets arise in various areas including coin weighing

*Dept. of Computer Science, University of Bonn. Email: hauptman@cs.uni-bonn.de

†Dept. of Computer Science, University of Bonn. Work supported by Hausdorff Doctoral Fellowship. Email: schmied@cs.uni-bonn.de

‡Dept. of Computer Science, University of Bonn. Work partially supported by Hausdorff Center for Mathematics, Bonn. Email: viehmann@cs.uni-bonn.de

problems [21], drug discovery [6], robot navigation [17], network discovery and verification [1], connected joins in graphs [18], and strategies for the Mastermind game [9]. The Metric Dimension problem has been widely investigated from the graph theoretical point of view [20, 6, 10, 3, 14, 22, 5, 4]. So far only a few papers discuss the computational complexity issues of this problem. The NP-hardness of the Metric Dimension problem was first mentioned in Gary and Johnson [11]. An explicit reduction from the 3-SAT problem was given by Khuller, Raghavachari, and Rosenfeld [17]. They also obtain for the Metric Dimension problem a $(2\ln(n) + \Theta(1))$-approximation algorithm based on the well-known greedy algorithm for the Set Cover problem and showed that the Metric Dimension problem is polynomial-time solvable for trees. Beerliova et al. [1] showed that the Metric Dimension problem (which they call the *Network Verification* problem) cannot be approximated within a factor of $o(\log(n))$ unless $P = NP$.

Berman, DasGupta, and Kao [2] study various *Test Set* problems and in particular give a $(1 + \ln(n))$-approximation algorithm for the *Test Set Collection* (TSC) problem. The Metric Dimension problem can be seen as a variant of the Test Set Collection problem where only certain combinations of tests (corresponding to the vertices of the input graph) are available (cf. Section 2.2).

Halldórsson, Halldórsson, and Ravi [12] study the Test Set Collection problem with bounded test size. They give a $(3 + 3\ln(k))$-approximation algorithm for the Test Set Collection problem with test of size at most k.

The approximation complexity of dense and superdense instances of various optimization problems was studied in Karpinski and Zelikovsky [16], see also Karpinski [15].

1.2 Our Contributions

This work is the first, best to our knowledge, providing explicit approximation lower bounds for both bounded degree and dense instances of the Metric Dimension problem. Furthermore, we improve the upper bounds for general and dense instances as well as the lower bound for general instances. We also observe that the Metric Dimension problem restricted to point sets in $\mathbb{R}^d$ is polynomial-time solvable whenever d is constant.

In particular, we prove that the Metric Dimension in graphs cannot be approximated to within a factor of $(1-\epsilon)\ln(n)$ for any constant $\epsilon > 0$, unless $NP \subset DTIME(n^{log(log(n))})$. Moreover, we give an $(1 + (1 + o(1))\ln(n))$-approximation algorithm based on a modified version of the approximation algorithm for the Test Set Collection problem from [2]. This improves the previously best approximation algorithm of Khuller, Raghavachari, and Rosenfeld with approximation ratio $(2\ln(n) + \Theta(1))$ [17].

For the Metric Dimension problem on bounded degree graphs, we prove that it is APX-hard with degree bound $B \geq 3$, and we provide explicit approximation lower bounds under the assumption $P \neq NP$.

By constructing an approximation preserving reduction from the Dominating Set problem on bounded degree graphs, we show that the Metric Dimension problem on k-superdense graphs is APX-hard for $k \geq 6$. We obtain explicit approximation lower bounds by combining this reduction with results from [8]. We also provide a constant-factor approximation algorithm with approximation ratio $(2 + 2\ln(k) + \ln(\log_2(k-1)) + o(1))$ for k-superdense instances.

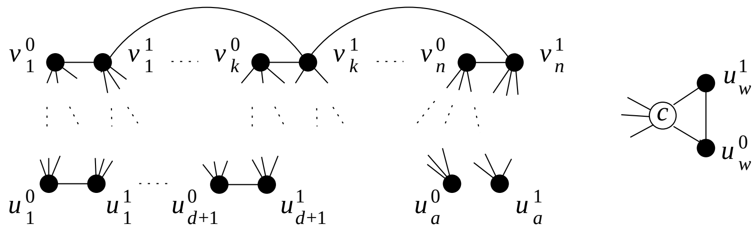

Figure 1: The graph $G' =: \tau_1(G)$

2 Metric Dimension of Graphs

In this section, we show that it is impossible (under reasonable complexity theoretic assumptions) to approximate the Metric Dimension of a Graph $G = (V, E)$ any better than $(1 - \epsilon)\ln(|V|)$ for any $\epsilon > 0$. We construct an approximation preserving reduction from the Dominating Set problem to the Metric Dimension problem.

2.1 Approximation Lower Bound

The Dominating Set problem is a special case of the Set Cover problem where we have to cover the vertex set of a given graph $G = (V, E)$ with sets from $\{N(v) \cup \{v\} \mid v \in V\}$. Here $N(v)$ denotes the set of neighbors of node v in G. Now we formulate our main result.

Theorem 2.1. *For any constant* $\epsilon > 0$*, the Metric Dimension problem cannot be approximated in polynomial time to within a factor of* $(1-\epsilon)\ln(n)$*, unless* $NP \subset DTIME(n^{log(log(n))})$.

Proof. In order to reduce the Dominating Set problem to the Metric Dimension problem, we have to convert a splitting problem into a covering problem. This will be done by introducing pairs of nodes for every element that needs to be covered. The pairs can only be distinguished by special vertices representing the sets $N(v) \cup \{v\}$.

The proof of the above theorem uses the following lemma.

Lemma 2.1. *There exists a polynomial-time computable function* τ_1 *that maps an instance* $G = (V, E)$ *of the Dominating Set problem into instance* $G' = (V', E')$ *of the Metric Dimension problem such that optimal solutions of* G *and* G'*,* OPT_{DS} *and* $dim_M(G')$ *respectively, satisfy the following:*

$$dim_M(G') \leq |OPT_{DS}| + \lceil \log_2(|V|) \rceil + 3$$

Proof. For notional simplicity, we introduce $n := |V|$ and $d := \lceil \log_2(n) \rceil$. The corresponding graph $G' = (V', E')$ contains for every $v_i \in V$ the pair of vertices $\{v_i^1, v_i^0\}$ and $2(d + 3)$ special vertices $u_1^k, ..., u_{d+1}^k, u_a^k, u_w^k$ with $k \in \{0, 1\}$. Finally, we add a vertex c which is connected to all other vertices. Furthermore, we connect u_k^1 and u_k^0 for all $k \in \{1, \dots, d + 1, w\}$. We join the vertices v_j^1 and v_j^0 with both u_k^1 and u_k^0 by an edge if and only if the binary representation of j has a 1 on the k-th position. The vertices u_a^1 and u_a^0 are both connected to all vertices v_i^j with $j \in \{0, 1\}$ and $v_i \in V$. Last of all, we add edges $\{v_i^1, v_j^1\}$ iff $\{v_i, v_j\} \in E$. The graph G' is depicted in Figure 1. In the following, we show that $B := \{u_1^1, ..., u_{d+1}^1, u_a^1, u_w^1\} \cup \{v_j^1 \mid v_j \in OPT_{DS}\}$ is a resolving set for G':
The pairs $\{c, v\}$ with $v \in V' \backslash \{c\}$ can be distinguished by u_a^1 and u_w^1. u_a^1 also resolves the pairs of the form $\{u_j^k, v_i^j\}$. In case of $\{u_j^0, u_{j'}^0\}$ with $a \neq j \neq j'$, we have $d(u_j^1, u_j^0) = 1$

and $d(u_j^1, u_{j'}^0) = 2$. Since the binary representation of numbers is unique, there is always a $u_j^1 \in B$ which can resolve $\{v_s^m, v_r^z\}$ with $s \neq r$. Now we are left with pairs of the form $\{v_j^1, v_j^0\}$. But these pairs are "covered" by v_l^1 with $v_l \in OPT_{DS}$.
Since the metric dimension of G' can be upper bounded by the cardinality of any resolving set, we conclude $dim_M(G') \leq |B| \leq d + 3 + |OPT_{DS}|$. □

The following lemma provides an algorithm that tranforms a solution for G' into a dominating set of the original graph G.

Lemma 2.2. *There is a polynomial-time computable function τ_2 that maps a solution B of $\tau_1(G)$ into solution DS of G such that $|DS| \leq |B|$ holds.*

Proof. At least one vertex of each pair u_i^1, u_i^0 with $\in \{1, .., d, a, w\}$ must be included in a resolving set B of G' since for all other vertices $v \in V' \backslash \{u_i^1, u_i^0\}$ we have $d(u_i^1, v) = d(v, u_i^0)$. Recall that this set can resolve any pair but p_i for all $v_i \in V$. Therefore, we have to determine which vertices are able to distinguish the remaining pairs. Notice that the only vertices which can resolve the pair p_i are exactly $w \in \{v_i^1, v_i^0, v_j^1 \mid v_j \in N(v_i)\}$. According to that fact, the set $DS(B) := \{v_k \mid \{v_k^1, v_k^0\} \cap B \neq \emptyset\}$ is a dominating set for G with $|DS(B)| \leq |B|$. □

In order to prove Theorem 2.1, we need the following straightforward extension of the hardness result in [8] for a restricted version of the Dominating Set problem.

Lemma 2.3. *Assuming $NP \not\subset DTIME(n^{log(log(n))})$, instances of the Dominating Set problem for which the optimal dominating set requires at least $log^2(n)$ vertices cannot be approximated to within a factor of $(1-\epsilon)\ln(n)$ for any $\epsilon > 0$ in polynomial time.*

We are ready to prove Theorem 2.1. Assume there exists a polynomial-time approximation algorithm $\mathcal{A}_1$ for the Metric Dimension problem with ratio $(1-\epsilon)\ln(n)$ for an $\epsilon > 0$. Next, we apply τ_1, $\mathcal{A}_1$, and τ_2 consecutively and get the following upper bound for the solution $DS(B)$ of the combined algorithms:

$$\begin{aligned} |DS(B)| &\leq |B| \leq (1-\epsilon)\ln(|V'|)\, dim_M(G') \leq (1-\epsilon)\ln(2n+2d+7)\, dim_M(G') \\ &\leq (1-\epsilon)\ln(n)\,[1+o(1)]\,(|OPT_{DS}(G)| + d + 3) \\ &\leq (1-\epsilon)\,[1+o(1)]\ln(n)|OPT_{DS}(DS)| \left[1 + \frac{d+3}{\Omega(\ln^2(n))}\right] \\ &\leq [1-\epsilon+o(1)]\ln(n)|OPT_{DS}(G)| \end{aligned}$$

This is a contradiction to Lemma 2.3, and Theorem 2.1 holds. □

2.2 Approximation Algorithm

In this section, we construct a $(1+(1+o(1))\ln(n))$-approximation algorithm for the Metric Dimension problem in graphs. This improves on the previous existing $(1+2\ln(n))$-approximation algorithm (cf. [17]). Our approximation algorithm is a variant of the algorithm of Berman, DasGupta, and Kao [2] for the TSC problem with an appropriately chosen information content function.

In the *TSC* problem, we are given an universe U and a subcollection of tests $T \subset P(U)$, and we ask for a set of tests $T' \subseteq T$ of minimum cardinality $|T'|$ such that $\forall e \in P_2(U)\ \exists t \in$

$T' : |t \cap e| = 1$. The following notations were introduced in [2]. A set of tests $T \subset S$ defines an equivalence relation $\equiv_T$ on U given by $[i \equiv j] \Leftrightarrow [\forall t \in T(i \in t \Leftrightarrow j \in t)]$. Let $A_1, .., A_k$ be the equivalence classes of $\equiv_T$, then the entropy of T is defined as $H_T = \log_2(\Pi_{i=1}^k |A_i|!)$. The information content of a test $t \in U$ with respect to T is defined as $IC(t, T) = H_T - H_{T \cup \{t\}}$. Berman, DasGupta, and Kao [2] provided the following simple greedy heuristic for the Min TSC problem:

Information Content Heuristic (ICH)

$T' := \emptyset$
while ($H_{T'} \neq 0$) do
  select a $t \in argmax_{t \in T \setminus T'}(IC(t, T'))$
  $T' := T' \cup \{t\}$
endwhile

Theorem 2.2. *([2]) ICH is a polynomial-time approximation algorithm for the TSC problem with ratio* $1 + \ln(\max_t IC(t, \emptyset))$.

We apply the ICH to the Metric Dimension problem, where the tests correspond to the vertices of G and a test splits the set V into possibly more than two classes of indistinguishable nodes. Hence, for each subset of $V' \subseteq V$ we have the associated equivalence relation $\equiv_{V'}$ given by:

$$u \equiv_{V'} w \iff [\forall v \in V' : d(v, u) = d(w, v)].$$

Modified ICH is now ICH applied to the information content function $IC(v, V') := H_{V'} - H_{V' \cup \{v\}}$.

Modified ICH

$V' := \emptyset$
while ($H_{V'} \neq 0$) do
  select a $v \in argmax_{v \in V \setminus V'}(IC(v, V'))$
  $V' := V' \cup \{v\}$
endwhile

Theorem 2.3. *Modified ICH is a polynomial-time approximation algorithm for the Metric Dimension problem with ratio* $1 + \ln(|V|) + \ln(\ln_2(|V|))$.

Proof. A test set corresponding to a vertex v in the Metric Dimension problem now partitions the vertex set V of G into at most n classes $A_0, .., A_{n-1}$ where $A_i := \{s \in U \mid dist(v, s) = i\}$. The procedure that partitions V into n classes can be thought of as a group of n tests each of which partitions V successively into 2 classes. Thus, we conclude $\max_{v \in V} IC(v, \emptyset) \leq \log_2(n!) - \log_2(1) \leq n \log_2(n)$. □

3 Metric Dimension of Bounded-Degree Graphs

In this section, we show that the Metric Dimension problem restricted to bounded degree graphs is APX-hard with degree bound $B \geq 3$. The Metric Dimension problem with degree bound $B \leq 2$ is in PO (see [17]). We construct an approximation preserving

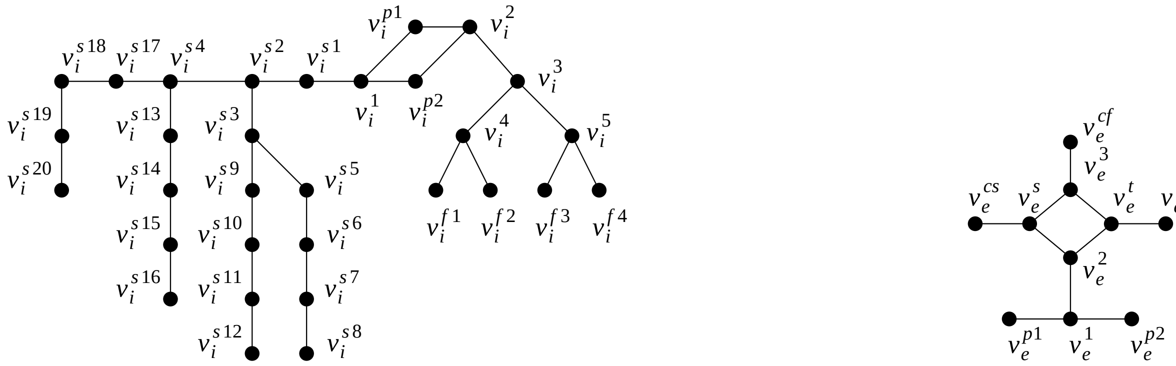

Figure 2: The graph G'_{v_i} is depicted in (a) and the graph G'_e in (b).

reduction from the bounded degree Vertex Cover problem and derive in this way the first explicit approximation lower bounds under the assumption $P \neq NP$. Now we formulate our theorem.

Theorem 3.1. *The B-bounded Metric Dimension problem is APX-hard for every $B \geq 3$ and is NP-hard to approximate within any constant better than $\frac{353}{352}$.*

Proof. Given a 4-regular graph G as an instance of Min-4-VC. We construct a graph G' similarly to the approximation preserving reduction in Theorem 2.1 and we introduce pairs of nodes representing the edges of the original graph that need to be covered. The high diameter of the graph G' is the main difficulty we have is to deal with. Since we cannot reach vertices quickly, we have to take care of pairs that are not supposed to be "covered".

The proof of the Theorem 3.1 uses the following lemma:

Lemma 3.1. *There exists a polynomial-time computable function f that maps every instance $G = (V, E)$ of MIN-4-VC to an instance $G' = (V', E')$ of the 3-Metric Dimension problem such that optimal solutions of $G = (V, E)$ and $G' = (V', E')$, VC^* and R^* respectively, satisfy $|R^*| \leq |VC^*| + |E| + |V|$.*

Proof. The graph G' consists of the the subgraphs G'_{v_i} for every $v_i \in V$ and G'_{e_j} for every $e_j \in E$. These subgraphs are depicted in Figure 2. We connect the vertex v_e^{cs} with exactly one $w \in \{v_i^8, v_i^{12}, v_i^{16}, v_i^{20}\}$ and v_e^{cf} with exactly one $x \in \{v_i^{f1}, v_i^{f2}, v_i^{f3}, v_i^{f4}\}$ if we have $v_i \in e_j$ and $deg_{G'}(x) = deg_{G'}(w) = 1$. The assignment of the vertices above is arbitrary as long as the degree of the vertices v_e^{cs} and v_e^{cf} is exactly 3 for every $e \in E$. See Figure 3 for an example of G'. Given a vertex cover VC of G, we show that $R := \{v_{e_j}^{p1}, v_i^{p1}, v_k^{s1} \mid v_i \in V, e_j \in E, v_k \in VC\}$ is resolving for G'.
First of all, we see that the set $R' := \{v_{e_j}^{p1}, v_i^{p1} \mid v_i \in V, e_j \in E\}$ can distinguish every pair of vertices of G' except the pairs $p_e^1 := \{v_e^{cs}, v_e^{ct}\}$ and $p_e^2 := \{v_j^s, v_j^t\}$ for every $e \in E$. Notice that the vertex v_k^{s1} can resolve the pairs p_e^1 and p_e^2 if and only if $v_k \in e$ holds. Since VC is a vertex cover for G, the set $R \backslash R'$ can distinguish the remaining pairs. Therefore, the metric dimension of G' can be bounded by $dim(G') \leq |R| = |E| + |V| + |VC|$. □

In order to construct an approximation preserving reduction, we need to transform a resolving set of G' into a vertex cover of G.

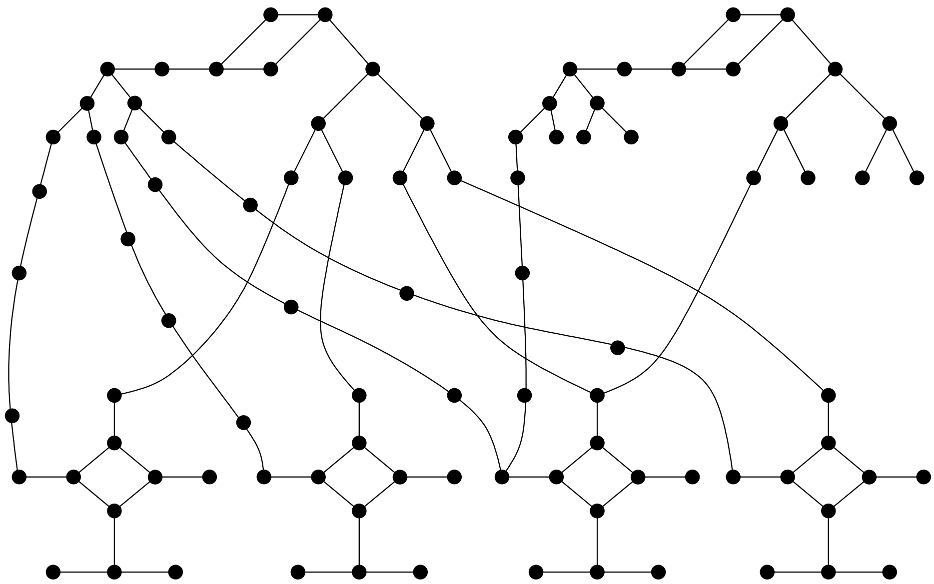

Figure 3: As an example, we illustrate a part of the graph G' which is constructed in the proof of Theorem 3.1/Lemma 3.1.

Lemma 3.2. *There exists a polynomial-time computable function f' that maps a resolving set R of $G' := f(G)$ into a vertex cover VC of G such that $|VC| \leq |R| - |E| - |V|$.*

Proof. In every resolving set, we have to have at least one vertex of $\{v_e^{p2}, v_e^{p1}\}$ for every $e \in E$ since these are the only vertices resolving themselves. The same holds for the pair $\{v_i^{p1}, v_i^{p2}\}$ for every $v_i \in V$. Recall that the set $\{v_e^{p1}, v_e^{p2}, v_i^{p1}, v_i^{p2} \mid e_j \in E, v_i \in V\}$ leaves only the pairs p_e^1 and p_e^2 for every $e \in E$ unresolved. Moreover, the only vertices which can distinguish both pairs p_e^1 and p_e^2 are $v \in p_e^1 \cup p_e^2 \cup \{v_i^{sj} \mid j \in \{1,..,17\}, i \in e\}$. Given a resolving set R of G', we build a set $VC(R)$ which is a vertex cover of G. For every $e := \{v_i, v_j\} \in E$, we add either v_i or v_j to $VC(B)$ iff $(p_e^1 \cup p_e^2) \cap R \neq \emptyset$. Further, we add v_i to $VC(R)$ iff $\{v_i^{sj} \mid j \in \{1,..,17\}\} \cap R \neq \emptyset$. Clearly, $VC(R)$ is a vertex cover of G with $|VC(R)| + |E| + |V| \leq |R|$. □

In order to prove Theorem 3.1, we use the following hardness result given in [7]:

Theorem 3.2. *([7]) Given a 4-regular graph $G = (V, E)$, let $OPT(G)$ denote the size of a minimal vertex cover of G. Then, the following partial decision problem is NP-hard to decide for $\epsilon \in (0, \frac{1}{2})$:*

$$|V|\frac{53 + 2\epsilon}{100} < OPT(G) \quad \textit{or} \quad |V|\frac{52 - 2\epsilon}{100} > OPT(G)$$

By applying Lemma 3.1 and 3.2, we construct a 3 bounded degree graph $G' = (V', E')$ for which the following is NP-hard to decide for all $\epsilon \in (0, \frac{1}{2})$:

$$|E| + |V| + |V|\frac{53 + 2\epsilon}{100} < dim(G') \quad \text{or} \quad |E| + |V| + |V|\frac{52 - 2\epsilon}{100} > dim(G')$$

Therefore, we conclude that it is NP-hard to approximate the Metric Dimension of a 3-bounded graph to within any constant better than $\frac{353}{352}$, which completes the proof of Theorem 3.1. □

4 Metric Dimension of k-Superdense Graphs

In this section, we study the approximation complexity of the Metric Dimension problem restricted to k-superdense graphs. We show that this special case is APX-hard to approximate for $k \geq 6$. In addition, we give an approximation algorithm with approximation ratio $(2 + 2\ln(k) + ln(log_2(k-1)) + o(1))$.

Note that the diameter of a k-superdense graph is at most 2. This fact will be crucial for our analysis.

4.1 Approximation Lower Bound

We construct an approximation preserving reduction from the *Dominating Set problem* on bounded degree graphs in order to obtain the following result. The explicit lower bounds will be given in Theorem 4.3.

Theorem 4.1. *The Metric Dimension problem on k-superdense graphs is APX-hard.*

Proof. For a given graph $G = (V, E)$ with vertex degree bound B, we construct a $(B+3)$-superdense graph G' in polynomial-time. G' consists of subgraphs G_v corresponding to every vertex v of G. Every resolving set of G' contains two vertices of every G_v that distinguish all pairs of vertices of G' except the special pairs $p_v = \{t_v, s_v\}$ of G_v for every $v \in V$. p_v can only be distinguished by vertices $u \in \{t_v\} \cup \{t_w \mid w \in N(v)\}$. This is equivalent to cover V with sets of the form $\{v\} \cup N(v)$.

Lemma 4.1. *There exists a polynomial-time computable function g that maps an instance $G = (V, E)$ with $\deg(v) \leq B$ for every vertex $v \in V$ of the Dominating Set problem into instance G' of the $(B+3)$-superdense Metric Dimension problem such that optimal solutions of G and G', OPT and OPT', respectively, satisfy the following: $OPT' \leq OPT + 2|V|$.*

Proof. For notational simplicity, we describe the complement graph $\overline{G'}$ of $G' = (V', E')$ given by $\overline{G'} := (V', \binom{V'}{2} \backslash E')$. The graph $\overline{G'}$ consists of the subgraphs G_i for every $v_i \in V$. The subgraph G_i is depicted in Figure 4(a). Finally, we connect the subgraphs G_i by adding edges $\{s_i, s_j\}$ iff $\{v_i, v_j\} \in E$.

Notice that the constructed graph G' is $(B+3)$-superdense. As an example, we constructed in Figure 4(b) a graph G and the corresponding complement graph of $g(G)$.

Let $DS(G)$ be an optimal dominating set of G. We claim that the set $R := \{a_i^1, b_i^1, s_j \mid v_i \in V, v_j \in DS(G)\}$ is resolving for G'. Firstly, we see that a_i^1 and b_i^1 can distinguish all pairs of vertices of the graph G_i except the pair $p_i := \{s_i, t_i\}$. Since we have $V \backslash N(a_i^1) \cup V \backslash N(b_i^1) = V(G_i)$ for all $v_i \in V$, $V \backslash N(a_i^1) \cap V \backslash N(a_j^1) = \emptyset$, and $V \backslash N(b_i^1) \cap V \backslash N(b_j^1) = \emptyset$ for $v_i \neq v_j$, the pairs of the form $\{x, y\}$ with $x \in V(G_i)$ and $y \in V(G_j)$ can be resolved.The remaining pairs p_i with $v_i \in V$ are distinguished by $\{s_j \mid v_j \in DS(G)\}$. Therefore, the metric dimension of G' can be bounded by $dim(G') \leq |R| = 2|V| + OPT$. □

Lemma 4.2. *There exists a polynomial-time computable function g' that maps every resolving set B of the graph $G' := g(G)$ into a dominating set $DS(B)$ of G such that $|DS(B)| \leq |B| - 2|V|$ holds.*

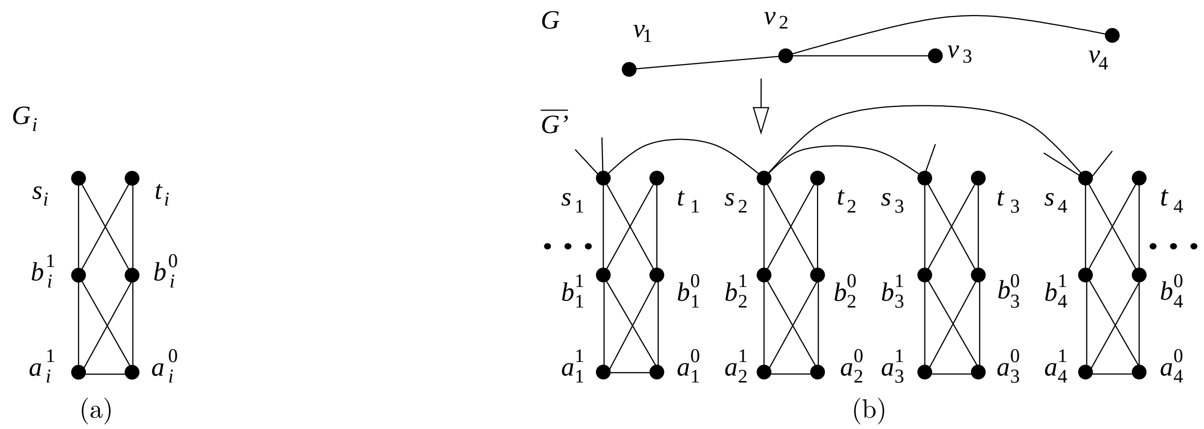

Figure 4: The graph G_i is depicted in (a). On the right-hand side the graph G as well as $\overline{G'}$ the corresponding complement of graph $g(G)$ are illustrated in (b).

Proof. In every resolving set B' of G' at least one of the vertices a_i^1, a_i^0 and at least one of b_i^1, b_i^0 for every $v_i \in V$ must be contained in B' since for every vertex u in $V(G')\backslash\{a_i^1, a_i^0\}$ we have $d(u, a_i^1) = d(u, a_i^0)$. An equivalent statement holds for b_i^1, b_i^0. Recall that the only pairs of vertices which cannot be distinguished by the set $\{a_i^1, b_i^1 \mid v_i \in V\}$ are $\{s_i, t_i\}$ for every $v_i \in V$. On the other hand, the only vertices which can distinguish the pair $\{s_i, t_i\}$ are given by $B(i) := \{s_i, t_i, s_j \mid v_i \in N(v_j)\}$. Therefore, every resolving set contains a $u \in B(i)$ for every $v_i \in V$. Clearly, $DS(B) := \{v_i \in V \mid B(i) \cap B \neq \emptyset\}$ is a dominating set of G with $|DS(B)| + 2|V| \leq |B|$. □

Since the Dominating Set Problem on bounded degree graphs is APX-hard, there exists a $r > 1$ such that it is NP-hard to approximate this problem with a better ratio than r. Assume, we could approximate the Metric Dimension on bounded degree graphs with a ratio $r' := \frac{r-1-\epsilon}{2(B+1.5)} + 1$ for any $\epsilon > 0$, then we get:

$$\begin{aligned}
|DS(B)| &\leq |B| - 2|V| \leq dim(G') \cdot r' - 2|V| \\
&\leq (2|V| + OPT(G)) \cdot r' - 2|V| = OPT(G)\left(r' + \frac{2(r'-1)|V|}{OPT(G)}\right) \\
&\leq OPT(G)\,[r' + 2(r'-1)(|B|+1)] \leq OPT(G) \cdot (r - \epsilon)
\end{aligned}$$

This is a contradiction to our assumption and Theorem 4.1 holds. □

Using the results of [8] for the B-Dominating Set problem, we get explicit approximation lower bounds. In particular, we make use of the following theorem:

Theorem 4.2. *[8] It is NP-hard to approximate the B-Dominating Set problem to within any constant better than $\frac{391}{390}$ for $B = 3$, $\frac{100}{99}$ for $B = 4$, and $\frac{53}{52}$ for $B = 5$.*

Combining this result with the approximation preserving reduction of Theorem 4.1, we obtain:

Theorem 4.3. *It is NP-hard to approximate the Metric Dimension on k-superdense graphs to within any constant better than $\frac{3511}{3510}$ for $k = 6$, $\frac{1090}{1089}$ for $k = 7$, and $\frac{677}{676}$ for $k = 8$.*

4.2 Approximation Algorithm

We combine the greedy approximation algorithm for the k-Set Cover problem with Modified ICH in order to obtain a $(2+2\ln(k)+\ln(\log_2(k-1))+o(1))$-approximation algorithm for the Metric Dimension problem in k-superdense graphs. Previously, Halldórsson et al. [12] used a similar approximation algorithm for the TSC problem with bounded test sizes, based on a twofold application of the greedy k-set cover algorithm. Here, we apply first the greedy k-set cover algorithm and afterwards use the Modified ICH to generate a resolving set. If we apply Modified ICH directly on an instance, in worst case we would only achieve an approximation ratio of $1 + ln(k) + \ln(\ln_2(n))$. So, we have to preprocess the vertex set by dividing it into small fractions at first in order to obtain a constant approximation ratio. Recall that in a k-superdense graph we have only $d(v, w) \in \{0, 1, 2\}$ and therefore only three equivalence classes occur. For every $v \in V$, let A_0^v, A_1^v and A_2^v be the equivalence classes under $\equiv_v$. Consider the following algorithm Pre-ICH:

1. Apply the greedy algorithm for the Min k-Set Cover problem to instance $SC(G) := (V, \{A_0^v \cup A_2^v \mid v \in V\})$ with solution $\{A_0^v \cup A_2^v \mid v \in V''\}$.
2. Apply Modified ICH with initial set $V' := V''$.

Theorem 4.4. *Pre-ICH is a* $(2+2\ln(k)+\ln(\log_2(k-1))+o(1))$*-approximation algorithm for the Metric Dimension problem on k-superdense graphs.*

Proof. Let G be a k-superdense graph and B the solution produced by Pre-ICH. In order to distinguish every pair in $P_2(V)$, every vertex but one must be contained in a set $A_0^w \cup A_2^w$, otherwise we would have two vertices u and x with $d(x, v) = d(u, v) = 1$ for every vertex v in a resolving set. Therefore, the optimal solution $OPT(SC(G))$ of $SC(G)$ can be upper bounded by $OPT(SC(G)) \leq dim(G) + 1$. Since the simple greedy heuristic for the k-Set Cover problem is a $(1 + \ln(k))$-approximation algorithm, we conclude $|V''| \leq (1 + \ln(k))(dim(G) + 1)$. Next, we want to derive an upper bound of the cardinality of $R := B\backslash V''$. We observe that the proof of Theorem 1 in [2] actually yields the following slightly more general result: When modified ICH is started with an initial set $T' := V''$ (instead of $T' := \emptyset$) it constructs a resolving set R of size $|R| \leq (1 + \ln(\max_{v \in V\backslash V''} IC(v, V\backslash V'')))dim(G)$. Thus, we have to analyze the worst-case behavior of the term $IC(v, V\backslash V'')$. Let $V'' = \{v_1, \ldots, v_c\}$ be the cover generated in step 1 of *Pre-ICH.* A node $v \in V \setminus V''$ might split an equivalence class into two or three parts, and furthermore $|A_0^v \cup A_2^v| \leq k$. Hence, v has a total budget of at most k to split classes, each of which is of size at most $k - 1$. When it splits a class of size s into two classes of size $s - a$ and a, this contributes $\log_2(\binom{s}{a})$ to the information content. This term is monotone in s, thus we have the following setup: We are given k classes each of size $k-1$ and another node v with budget k, and we ask for an upper bound of the information content when v spends this budget into splitting the classes into two or three parts (at most one is split into three). Consider two classes that are split by v by use of budget a and b respectively, where $1 \leq b \leq a < \frac{k-1}{2}$. From $\binom{k-1}{a}\binom{k-1}{b} \geq \binom{k-1}{a+1}\binom{k-1}{b-1}$ we see that the contribution to the information content is maximized when $a = 1$. Now we consider the case that v splits a class into three parts of size $1, x$ and $k - 1 - x - 1$ respectively, where $x \leq \frac{k-1}{2}$. Since the contribution to the information content is $\log_2[(k-1) \cdot \binom{k-2}{x}]$, by the same argument as before we get that this term attains a maximum for $x = 1$. Therefore, an upper bound for the maximum is attained when v splits k classes, each by using an amount of 1 from its budget. For notational simplicity, we set $c := \lceil n/k \rceil$ and get

$$IC(v, V \backslash V'') \leq \log_2 \left(\frac{(1!)^c [(k-1)!]^c}{(1!)^{(c+k)} [(k-2)!]^k [(k-1)!]^{(c-k)}} \right) \leq k \log_2(k-1).$$

We are ready to analyze the approximation ratio of Pre-ICH:

$$\begin{aligned} |B| &\leq |R| + |V''| \\ &\leq [1 + \ln(k) + 1 + \ln(k) + \ln(\log_2(k-1))]\, dim(G) + 1 + \ln(k) \\ &\leq [2 + 2\ln(k) + \ln(\log_2(k-1)) + o(1)] dim(G) \end{aligned}$$

In the last inequality, we used the facts $\frac{n}{k} \leq dim(G) + 1$ and $k = \Theta(1)$. □

5 Metric Dimension in $\mathbb{R}^d$

In this section, we consider the Metric Dimension problem in $\mathbb{R}^d$.

Theorem 5.1. *For each $d \in \mathbb{N}$, the Metric Dimension problem restricted to finite sets of points in $\mathbb{R}^d$ with the Euclidean distance is in PO.*

Proof. Let $X = \{x_1, \ldots, x_n\} \subset \mathbb{R}^d$ and assume X spans the $\mathbb{R}^d$ - otherwise we replace $\mathbb{R}^d$ by the subspace generated by X. Let d_2 denote the Euclidean distance in $\mathbb{R}^d$ and $\langle \cdot, \cdot \rangle$ the inner product. For $1 \leq i < j \leq n$, the set of points which cannot distinguish x_i and x_j is an affine hyperplane

$$I_{ij} = \{x \in \mathbb{R}^d | d_2(x, x_i) = d_2(x, x_j)\} = \{x \in \mathbb{R}^d | \langle x - m_{ij}, x_i - x_j \rangle = 0\}$$

with $m_{ij} = \frac{1}{2}(x_i + x_j)$. Consider a set $X' = \{x_{i_0}, \ldots, x_{i_d}\} \subseteq X$ such that the $x_{i_j} - x_{i_0}, j = 1, \ldots d$ are linearly independent. Assume X' is not a resolving set for X, then $X' \subset I_{ij}$ for some $1 \leq i < j \leq n$ which would be a contradiction to the $x_{i_j} - x_{i_0}$ being linearly independent. Thus, we can construct a minimum cardinality resolving set by enumerating all subsets of X of size at most $d + 1$. □

Acknowledgment

We thank Marek Karpinski for a number of interesting discussions and for his support.

References

[1] Z. Beerliova, F. Eberhard, T. Erlebach, A. Hall, M. Hoffmann, M. Mihalák, and L. Ram, *Network Discovery and Verification*, IEEE J. on Selected Areas in Communications 24, pp. 2168–2181, 2006.

[2] P. Berman, B. DasGupta, and M. Kao, *Tight Approximability Results for Test Set Problems in Bioinformatics*, J. Comput. Syst. Sci. 71, pp. 145–162, 2005.

[3] J. Cáceres, C. Hernando, M. Mora, I. Pelayo, M. Puertas, C. Seara, and D. Wood, *On the Metric Dimension of Cartesian Products of Graphs*, SIAM J. Disc. Math. 21, pp. 423–441, 2007.

[4] J. Cáceres, C. Hernando, M. Mora, I. Pelayo, and M. Puertas, *On the Metric Dimension of Infinite Graphs*, Electr. Notes in Disc. Math. 35, pp. 15–20, 2009.

[5] G. Chappell, J. Gimbel, and C. Hartman, *Bounds on the Metric and Partition Dimensions of a Graph*, Ars Combinatoria 88, pp. 349–366, 2008.

[6] G. Chartrand, L. Eroh, M. Johnson, and O. Oellermann, *Resolvability in Graphs and the Metric Dimension of a Graph*, Disc. Appl. Math. 105, pp. 99–133, 2000.

[7] M. Chlebík and J. Chlebíková, *Complexity of Approximating Bounded Variants of Optimization Problems*, Theor. Comput. Sci. 354, pp. 320–338, 2006.

[8] M. Chlebík and J. Chlebíková, *Approximation Hardness of Dominating Set Problems in Bounded Degree Graphs*, Inf. Comput. 206, pp. 1264–1275, 2008.

[9] V. Chvátal, *Mastermind*, Combinatorica 3, pp. 325–329, 1983.

[10] M. Fehr, S. Gosselin, and O. Oellermann, *The Metric Dimension of Cayley digraphs*, Disc. Math. 306, pp. 31–41, 2006.

[11] M. Garey and D. Johnson, *Computers and Intractability: A Guide to the Theory on NP-Completeness*, Freeman, New York, 1979.

[12] B. Halldórsson, M. Halldórsson, and R. Ravi, *On the Approximability of the Minimum Test Collection Problem*, Proc. 9th ESA (2001), LNCS 2161, pp. 158–169.

[13] F. Harary and R. Melter, *On the Metric Dimension of a Graph*, Ars Combinatoria 2, pp. 191–195, 1976.

[14] C. Hernando, M. Mora, I. Pelayo, C. Seara, and D. Wood, *Extremal Graph Theory for Metric Dimension and Diameter*, Electr. Notes in Disc. Math. 29, pp. 339–343, 2007.

[15] M. Karpinski and W. Schudy, *Linear Time Approximation Schemes for the Gale-Berlekamp Game and Related Minimization Problems*, Proc. 41st ACM STOC (2009), pp. 313–322.

[16] M. Karpinski and A. Zelikovsky, *Approximating Dense Cases of Covering Problems*, Proc. DIMACS Workshop on Network Design: Connectivity and Facilities Location (1997), pp. 169–178; also published in ECCC TR97-004, 1997.

[17] S. Khuller, B. Raghavachari, and A. Rosenfeld, *Landmarks in Graphs*, Disc. Appl. Math. 70, pp. 217–229, 1996.

[18] A. Sebö and E. Tannier, *On Metric Generators of Graphs*, Math. Oper. Res. 29, pp. 383–393, 2004.

[19] P. Slater, *Leaves of Trees*, Southeastern Conf. on Combin., Graph Theory, and Comp. (1975), Congressus Numerantium 14, pp. 549–559.

[20] P. Slater, *Dominating and Reference Sets in Graphs*, J. Math. Phys. Sci. 22, pp. 445–455, 1988.

[21] H. Shapiro and S. Söderberg, *A Combinatory Detection Problem*, Amer. Math. Monthly 70, pp. 1066-1070, 1963.

[22] I. Tomescu, *Discrepancies between Metric Dimension and Partition Dimension of a Connected Graph*, Disc. Math. 308, pp. 5026–5031, 2008.