On the approximation hardness of geodetic set and its variants

Tom Davot, Lucas Isenmann, Jocelyn Thiebaut

On the approximation hardness of geodetic set and its variants

Tom Davot[1], Lucas Isenmann[2], and Jocelyn Thiebaut[3]

[1] LIRMM, Univ Montpellier, CNRS, Montpellier, France davot@lirmm.fr
[2] Aix-Marseille Université, France lucas.isenmann@laposte.net
[3] INSA Centre Val de Loire, Univ. Orléans, LIFO EA 4022, Orléans, France jocelyn.thiebaut@univ-orleans.fr

Abstract. Given a graph, a geodetic set (resp. edge geodetic set) is a subset of its vertices such that every vertex (resp. edge) of the graph is on a shortest path between two vertices of the subset. A strong geodetic set is a subset S of vertices and a choice of a shortest path for every pair of vertices of S such that every vertex is on one of these shortest paths. The geodetic number (resp. edge geodetic number) of a graph is the minimum size of a geodetic set (resp. edge geodetic set) and the strong geodetic number is the minimum size of a strong geodetic set. We first prove that, given a subset of vertices, it is $\mathcal{NP}$-hard to determine whether it is a strong geodesic set. Therefore, it seems natural to study the problem of maximizing the number of covered vertices by a choice of a shortest path for every pair of a provided subset of vertices. We provide a tight 2-approximation algorithm to solve this problem. Then, we show that there is no $^{781}/_{780}$ polynomial-time approximation algorithm for edge geodetic number and strong geodetic number on subcubic bipartite graphs with arbitrarily high girth. We also prove that geodetic number and edge geodetic number are both LOG-$\mathcal{APX}$-hard, even on subcubic bipartite graphs with arbitrarily high girth. Finally, we disprove a conjecture of Iršič and Konvalinka by proving that the strong geodetic number can be computed in polynomial time in complete multipartite graphs.

1 Introduction

Geodetic number and edge geodetic number. A *geodesic* between two vertices of a graph G is a path of minimum length between x and y. The *geodetic number* of G is the minimum size of a subset X of the vertices such that, for every vertex v, there exists a geodesic between two vertices of X containing v. The geodetic number of a graph has been introduced by Harary et al in [13], where the authors showed that deciding whether a graph has a geodetic number less than an integer k is $\mathcal{NP}$-complete. The complexity of this problem has also been investigated in several classes of graphs, such as bipartite graphs [11] and chordal graphs [10]

Preprint Version
The final authentificated version is available online at https://doi.org/10.1007/978-3-030-89543-3_7

where it remains $\mathcal{NP}$-complete. Recently, Chakraborty *et al.* proved that finding the geodetic number of a graph is $\mathcal{NP}$-hard on planar graphs with maximum degree six and line graphs [6]. They also proved in [7] that, unless $\mathcal{P} = \mathcal{NP}$, there is no polynomial time $o(\log n)$-approximation algorithm for computing the geodetic number of a graph, even on graphs that have a universal vertex and where n stands for the number of vertices in the input graph.

The edge version of the geodetic number has been introduced independently in [3] and in [22]. A subset X of the vertices is an *edge geodetic set* if, for every edge e, there is a geodesic between two vertices of X containing e. The *edge geodetic number* of G is the size of the smallest geodetic set of G. Note that, given a graph G, the geodetic number of G is smaller than its edge geodetic number. This edge version is also known to be $\mathcal{NP}$-hard [3]. This problem has been studied on several classes of graphs, such as Cartesian products [1, 21] and fuzzy graphs [20]. From a structural point of view, Santhakumaran and Ullas Chandran characterized graphs with a prescribed edge geodetic number [23]. For more results and motivations about geodetic sets, see [5].

Strong geodetic number and strong edge geodetic number. A subset of vertices X is a *strong geodetic set* if there exists a function $\tilde{I}$ that associates a unique geodesic to each pair of vertices of X and such that every vertex v is contained in a geodesic $\tilde{I}(a, b)$, where $\{a, b\} \subseteq X$. In the following, we call such a function a *geodesic assignation* for X. The *strong geodetic number* of a graph G is the size of the smallest strong geodetic set. The strong geodetic number has been introduced recently by Arokiaraj *et al.* [2]. In their original paper, the authors motivate this variation by social network applications. Furthermore, they also prove that this problem is $\mathcal{NP}$-complete. Note that it remains $\mathcal{NP}$-hard even when restricted to bipartite graphs [15]. This problem has been studied on complete Apollonian networks [2], grids and cylinders [16], and on Cartesian product of graphs [12], on complete bipartite graphs [14], on complete multipartite graphs [15] and on outerplanar graphs [18]. Connections to the diameter of the graph were studied in [14] and to the isometric path problem [2].

Finally, the edge version of the strong geodetic problem, where we want to cover every edge of the graph, has been introduced by Manuel *et al.* [17] and were proved $\mathcal{NP}$-complete. Zmazek recently studied the values of the edge strong geodetic number on grids [25].

Our results. The results of our paper are divided in four sections. In Section 3, we propose a variant of the strong geodetic problem where, given a subset S of vertices, the question is to determine whether S is a strong geodesic set of the graph. Using a reduction from Monotone Balanced 3-SAT-(4), we prove that this problem is $\mathcal{NP}$-hard. Then, we consider in Section 4 the problem of maximizing the number of covered vertices by a choice of a geodesic for each pair of a provided subset of vertices, and provide a tight 2-approximation algorithm to solve it. In Section 5, we reduce the geodetic problems from Set Cover. We first give it in the general case, and then we adapt the previous construction on bipartite graphs with arbitrarily high girth. Using the previous reductions, we show in Section 6 that there is no approximation of Edge Geodetic Number

with an approximation factor better than $^{781}/_{780}$. We also prove that geodetic number and edge geodetic number are both LOG-$\mathcal{APX}$-hard, even on subcubic bipartite graphs with arbitrarily high girth. Finally, in section 7, we give a polynomial time algorithm which computes the Strong Geodetic Number of complete multipartite graphs, disproving the conjecture of [15] which states that Strong Geodetic Number is $\mathcal{NP}$-hard on complete multipartite graphs.
Due to space constraints, the proofs have been omitted. However, a full version can be found in `https://hal-lirmm.ccsd.cnrs.fr/lirmm-03328636`.

2 Notations

We first introduce some notations and formally define the problems. Given a set X, we denote by $\mathcal{P}_2(X)$ the set of its pairs. Given two sets X and Y, we denote by $X \sqcup Y$ the union $X \cup Y$ when X and Y are disjoint. Let G be a graph, we denote by $V(G)$ its set of vertices and by $E(G)$ its set of edges. We denote by $D_1(G)$ the set of vertices of degree one in G. Let X be a subset of vertices of G and x be a vertex, we say that x is *selected* by X if $x \in X$ and that x is *covered* by X if x is contained in a geodesic between two vertices of X (or simply selected or covered if there is no ambiguity on X). Likewise, let uv be an edge, we say that uv is *covered* by X (or simply covered) if uv is contained in a geodesic between two vertices of X.

To describe a path between two vertices u and v, we introduce the operator $\sim$ as follows. Let $p_1, p_2, \dots, p_k$ be some subgraphs such that $u \in p_1, v \in p_k$ and for each $i < k$, there exists a unique vertex $x_i \in p_i \cap p_{i+1}$. The path described by $p_1 \sim p_2 \sim \cdots \sim p_k$ corresponds to $(u, \dots, x_1, \dots, x_i, \dots, x_{k-1}, \dots, v)$ (following successively the paths $p_1, p_2, \dots, p_k$). Let g be a path that contains the vertices u and v. We denote by $g[u, v]$ the subpath of g with extremities u and v. Furthermore, we denote by $V(g)$ the vertices of g. Similarly, given a geodesic assignation $\tilde{I}$ for a set of vertices X, we denote by $V(\tilde{I})$ the vertices covered by the geodesics of $\tilde{I}$.
We now introduce the problems studied in this work.

(Strong) (Edge) Geodetic Number
Input: a simple graph G and an integer k.
Question: is there a (strong) (edge) geodetic set $X \subseteq V$ of size k?

The following already known property will be fundamental in the proofs of our reductions as it helps to force some vertices to be part of a (strong) (edge) geodetic set.

Property 1. If G is a graph and X is a solution of any geodesic problem, then we have $D_1(G) \subseteq X$.

3 Hardness to find a geodesic assignation

In the proof that computing the strong geodesic number is $\mathcal{NP}$-complete, the geodesic assignation is rather trivial [2]. In this section, we show that determining if a set of vertices is a strong geodetic set (*i.e.* computing a geodesic assignation) is in itself $\mathcal{NP}$-complete. To do so, we reduce from a special case of 3-SAT called MONOTONE BALANCED 3-SAT-(4). In this variant, the boolean formula is composed of *monotone* clauses, that is, clauses that contains only positive literals or only negative literals. MONOTONE BALANCED 3-SAT-(4) is defined as follows.

MONOTONE BALANCED 3-SAT-(4)
Input: a monotone 3-SAT formula φ where each variable occurs exactly two times positively and two times negatively.
Question: is φ satisfiable?

Darman and Döcker showed that this problem is $\mathcal{NP}$-complete [8]. We introduce the following construction.

Construction 1. *Let φ be a* MONOTONE BALANCED 3-SAT-(4) *formula, we construct the following graph G:*
- *For each clause C_j, introduce a vertex q_j.*
- *For each variable x_i, introduce two edges $v_i^0v_i^1$ and $u_i^0u_i^1$. Furthermore, let C_j and $C_{j'}$, with $j < j'$ be the two clauses where x_i occurs with the same polarity (i.e. it appears positively in both clauses, or negatively in both), construct a path $(v_i^1, q_j, q_{j'}, u_i^1)$.*
- *For each pair of vertices v_i^1 and $u_{i'}^1$ with $i \neq i'$, introduce a vertex $t_{i,i'}$ and construct the path $(v_i^1, t_{i,i'}, u_{i'}^1)$.*
- *Finally, construct two vertices k_v and k_u, and for each variable x_i, introduce the edges $v_i^1k_v$ and $u_i^1k_u$.*

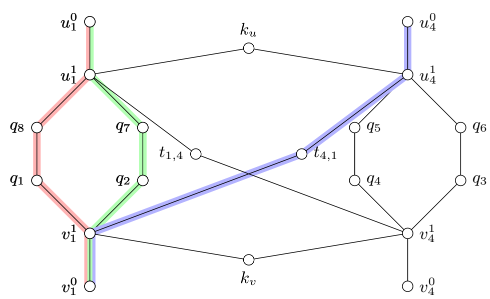

Fig. 1. Example of a subgraph induced by Construction 1. In the Boolean formula, the variable x_1 appears positively in C_2 and C_7 and negatively in C_1 and C_8. The variable x_4 appears positively in C_3 and C_6 and negatively in C_4 and C_5. The paths p_1, $\bar{p}_1$ and $p_{1,4}$ are depicted in green, red and blue, respectively.

For each variable x_i, let C_j and $C_{j'}$ (resp. C_k and $C_{k'}$) with $j < j'$ be the clauses where x_i occurs positively (resp. negatively). We denote by p_i the path $(v_i^0, v_i^1, q_j, q_{j'}, u_i^1, u_i^0)$, by $\bar{p}_i$ the path $(v_i^0, v_i^1, q_k, q_{k'}, u_i^1, u_i^0)$ and by $p_{i,i'}$ the path $(v_i^0, v_i^1, t_{i,i'}, u_{i'}^1, u_{i'}^0)$, for any $i \neq i'$. An example of a graph produced by Construction 1 is depicted in Figure 1.

Lemma 1. *Let φ be a* MONOTONE BALANCED 3-SAT-(4) *formula and G its graph resulting from Construction 1. Let $\tilde{I}$ be a geodesic assignation for $D_1(G)$.*

It is possible to construct a geodesic assignation $\tilde{I}'$ for $D_1(G)$ such that $|V(\tilde{I}')| \leq |V(\tilde{I})|$, and:
(1) for any $i \neq i'$, the geodesic between v_i^0 and $u_{i'}^0$ in $\tilde{I}'$ is $p_{i,i'}$, and
(2) for any i, the geodesic between v_i^0 and u_i^0 in $\tilde{I}'$ is either p_i or $\bar{p}_i$.

Proof. First, note that every geodesic with extremity v_i^0 (resp. u_i^0) covers the vertex v_i^1 (resp. u_i^1). Second, the geodesic between two vertices u_i^0 and $u_{i'}^0$ is $(u_i^0, u_i^1, k_u, u_{i'}^1, u_{i'}^0)$. Similarly, the geodesic between two vertices v_i^0 and $v_{i'}^0$ is $(v_i^0, v_i^1, k_v, v_{i'}^1, v_{i'}^0)$. Hence, it remains to cover the q_j vertices and the $t_{i,i'}$ vertices. Let $V_1 = \{v_i^1 \mid x_i \in \varphi\}$ and $U_1 = \{u_i^1 \mid x_i \in \varphi\}$. Since there is no edge between V_1 and U_1, the distance between any vertex of V_1 and any vertex of U_1 is at least two. From that, we can deduce that no geodesic g between two vertices v_i^0 and $u_{i'}^0$ contains the vertex k_v or the vertex k_u since the length of g would be at least six, and $p_{i,i'}$ or p_i (if $i = i'$) are shorter paths. Let us now prove the two items of the statement.

(1) Since the distance between V_1 and U_1 is at least two, the path $p_{i,i'}$ is a geodesic between v_i^0 and $u_{i'}^0$. Suppose that there exists another geodesic g between v_i^0 and $u_{i'}^0$. We have $g = (v_i^0, v_i^1, q_j, u_{i'}^1, u_{i'}^0)$, where the clause C_j contains x_i and $x_{i'}$ in φ. Suppose, without loss of generality that C_j contains positive literals. Toward a contradiction, suppose that $t_{i,i'}$ is covered by $\tilde{I}$ and let g' be the geodesic that covers $t_{i,i'}$. Since g' cannot have the same extremities as g, g' contains either the vertex k_v or the vertex k_u, contradicting g' being a geodesic. Hence, the vertex $t_{i,i'}$ is not covered by $\tilde{I}$, then we can replace g by $p_{i,i'}$ in it to obtain a solution as thought.

(2) Let g be a geodesic between v_i^0 and u_i^0. Since, g does not contain k_v, then v_i^1 is adjacent to a vertex q_j in g. Likewise, u_i^1 is adjacent to a vertex $q_{j'}$ in g. By contradiction, suppose that x_i appears positively in C_j and negatively in $C_{j'}$. Then, since the edge $q_j q_{j'}$ exists, by construction, there is a path $p_{i'}$ or $\bar{p}_{i'}$ containing q_j and $q_{j'}$. But then, the variable $x_{i'}$ appears either positively in C_j and $C_{j'}$ or negatively in C_j and C_j'. In any case, we reach a contradiction since either C_j or $C_{j'}$ is not monotone. Hence, x_i appears only positively in C_j and $C_{j'}$ or only negatively in C_j and $C_{j'}$, and then g corresponds to either p_i or $\bar{p}_i$.

Theorem 1. *It is $\mathcal{NP}$-hard to determine if a set of vertices V' is a strong geodetic set even if, for every strong geodetic set V_{strong}, we have $V' \subseteq V_{strong}$.*

Proof. Let φ be a Monotone Balanced 3-SAT-(4) formula and G its graph resulting from Construction 1. We show that φ is satisfiable if and only if $D_1(G)$ is a strong geodetic set.

- Let β be a satisfying assignment of φ, we construct a geodesic assignation $\tilde{I}$ for $D_1(G)$ as follows. For each $i \neq i'$, we set $\tilde{I}(v_i^0, u_i^0) = p_{i,i'}$, $\tilde{I}(v_{i'}^0, u_i^0) = p_{i',i}$, $\tilde{I}(v_i^0, v_{i'}^0) = (v_i^0, k_v, v_{i'}^0)$, and $\tilde{I}(u_i^0, u_{i'}^0) = (u_i^0, k_u, u_{i'}^0)$. Hence, every vertex is covered except the q_j vertices. Now, for each variable x_i, if x_i is assigned to *true*, we set $\tilde{I}(v_i^0, u_i^0) = p_i$ and we set $\tilde{I}(v_i^0, u_i^0) = \bar{p}_i$, otherwise.

Suppose, there is a vertex q_j that is not covered by $\tilde{I}$. Let x_i, x_k and x_ℓ be the three variables that occur in C_j and suppose that C_j contains positive literals. We have $img(\tilde{I}) \cap \{p_i, p_k, p_\ell\} = \emptyset$, and then $\beta(x_i) = \beta(x_k) = \beta(x_\ell) = false$, contradicting β being a satisfying assignment for φ. Hence, every vertex of G is covered and $D_1(G)$ is a strong geodetic set of G.

– Let $\tilde{I}$ be a geodesic assignation for $D_1(G)$ that covers every vertex of G, and such that $\tilde{I}$ respects properties of Lemma 1. We construct a satisfying assignment β for φ as follows. For each variable x_i, if $p_i \in img(\tilde{I})$, we set $\beta(x_i) = true$ and $\beta(x_i) = false$, otherwise.
Suppose there is a clause C_j that is not satisfied by β and suppose by symmetry that C_j contains positive literals. Let x_i, x_k and x_ℓ be the three literals of C_j. We have $img(\tilde{I}) \cap \{p_i, p_k, p_\ell\} = \emptyset$ and then q_j is not covered by $\tilde{I}$ which is a contradiction. Hence, β is a satisfying assignment for φ.

Finally, by Property 1, the set $D_1(G)$ belongs to any strong geodetic set.

From this theorem, we can derive a result about the residue variant of Strong Geodetic Number. The *residue variant* of an optimisation problem has been defined recently in [24] and consists of, given a partial solution P for an instance I, finding an optimal partial solution R such that $P \cup R$ is a solution for I. The complexity class $\mathcal{RAPX}$ contains the residue variant optimisation problems such that the score of the residue can be approximated by a constant.

Corollary 1. Strong Geodetic Number $\notin \mathcal{RAPX}$

Proof. Let φ be a Monotone Balanced 3-SAT-(4) formula and G be its graph resulting from Construction 1. Then, given a partial solution $P = D_1(G)$, G has a residue solution $R = \emptyset$ if and only if φ is satisfiable. Hence, the residual variant of Strong Geodetic Number cannot be approximated to any constant factor unless $\mathcal{P} = \mathcal{NP}$.

4 Approximation

Since it is hard to determine if a subset of vertices is a strong geodetic set, a natural question that arises is to find, given a subset of vertices, a geodetic assignation that maximizes the number of covered vertices. We call this problem Max Geodesic Assignation. By Theorem 1, this problem is also $\mathcal{NP}$-hard and we show that this problem belongs to $\mathcal{APX}$, *i.e.* approximable within a constant ratio. In this part, we show that this problem is 2-approximable using a simple greedy algorithm, defined in Algorithm 1.

Theorem 2. *Algorithm 1 computes in polynomial time a solution for* Max Geodesic Assignation *with an approximation ratio of* 2 *and this ratio is tight.*

Proof. Let $\tilde{I}_{app}$ be the geodesic assignation computed by Algorithm 1 and let $\tilde{I}_{opt}$ be an optimal geodesic assignation. We show that there exists an application $f : V(\tilde{I}_{opt}) \to V(\tilde{I}_{app})$ such that for each $u \in V(\tilde{I}_{app})$, $|f^{-1}(u)| \leq 2$.

Algorithm 1: Greedy Algorithm

Data: A graph G and a set of vertices $V' \subseteq V(G)$.
Result: A geodetic assignation $\tilde{I}$ for V'.

$A \leftarrow \mathcal{P}_2(V')$;
while $A \neq \emptyset$ **do**
  $\{u, v\} \leftarrow$ first element of A;
  $g \leftarrow$ geodesic between u and v that maximizes $|V(g) \setminus V(\tilde{I})|$;
  Set $\tilde{I}(u, v) := g$; $A \leftarrow A \setminus \{\{u, v\}\}$;
end
return $\tilde{I}$;

First, for each vertex $v \in V(\tilde{I}_{opt}) \cap V(\tilde{I}_{app})$, we set $f(v) = v$. Further, let g^i_{app} the geodesic chosen by the greedy algorithm at step i and let g^i_{opt} be the geodesic with the same extremities in $\tilde{I}_{opt}$. For each i, let $V^i_{opt} = V(g^i_{opt}) \setminus \bigcup_{j<i} V(g^j_{opt})$ and $V^i_{app} = V(g^i_{app}) \setminus \bigcup_{j<i} V(g^j_{app})$ (*i.e.* the set of vertices newly covered by g^i_{app}). We have $|V^i_{opt} \setminus V(\tilde{I}_{app})| \leq |V^i_{app}|$ since otherwise, the greedy algorithm would have chosen g^i_{opt} at step i. Thus, there exists an injective function $f' : V^i_{opt} \setminus V(\tilde{I}_{app}) \to V^i_{app}$ and, for each vertex $v \in V^i_{opt} \setminus V(\tilde{I}_{app})$, we set $f(v) = f'(v)$. Since each vertex $u \in V(\tilde{I}_{app})$ belongs to a unique V^i_{app}, we have $|f^{-1}(u)| \leq 2$. Moreover, since each vertex $v \in V(\tilde{I}_{opt})$ belongs either to $V(\tilde{I}_{opt}) \cap V(\tilde{I}_{app})$ or to a set $V^i_{opt} \setminus V(\tilde{I}_{app})$, we defined a function f as thought. It follows that $|V(\tilde{I}_{opt})| \leq 2 \cdot |V(\tilde{I}_{app})|$, proving the approximation ratio. Furthermore, the ratio is tight, as shown by Figure 2.

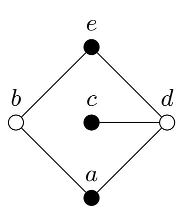

Fig. 2. Tightness of the approximation ratio of Algorithm 1. Consider a, c and e as selected (in black in the graph). The optimal solution consists in taking the geodesics (a, b, e), (c, d, e) and (a, c, d) which cover the non-selected vertices b and d. The greedy algorithm can start by taking the geodesic (a, d, e) between a and e. Then the algorithm will choose (c, d, e) and (a, c, d) for the last two pairs. This leads to a set of geodesics which only covers d.

5 Reduction from Set Cover

In this part, we prove preliminary results that will be used in the next section. More specifically, we reduce the geodetic problems from the classic $\mathcal{NP}$-complete problem Set Cover described as follows.

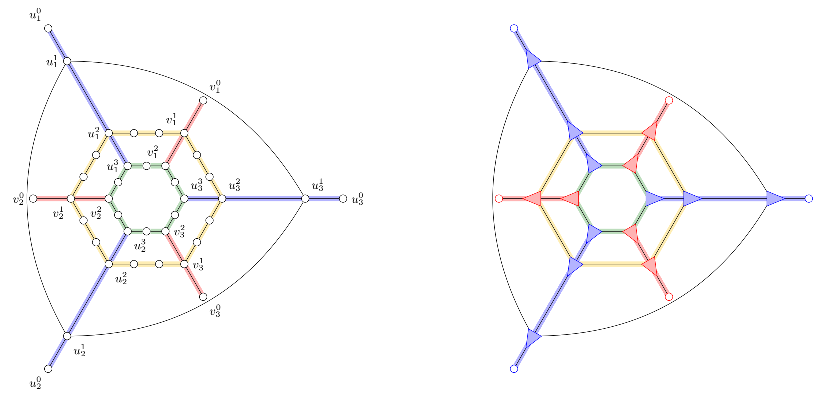

Fig. 3. Example of a graph produced by Construction 2 (left) and Construction 4 (right) on the collection containing $S_1 = \{E_1, E_3\}, S_2 = \{E_1, E_2\}$ and $S_3 = \{E_2, E_3\}$. Element paths, set paths, cut paths and long paths are coloured in blue, red, yellow and green, respectively. In the right graph, black edges, green edges and yellow edges represent paths of length $h, 2h$ and $3h$, respectively

SET COVER (SC)
Input: A collection $C = \{S_1, \ldots, S_{m'}\}$ of finite sets over the universe $U = \{E_1, \ldots, E_{n'}\}$.
Question: Find a minimum $C' \subseteq C$ such that every element of U is contained in a set of C'.

For the strong versions, we use a version of SET COVER, denoted (k, k')-SET COVER, where the size of the intersection between two sets is at most k and the set sizes are bounded by k'. Notice that since VERTEX COVER is a particular case of $(1, k')$-SET COVER, then (k, k')-SET COVER is $\mathcal{NP}$-complete.
In the following, we first show how this reduction works in the general case and then, we adapt it in subcubic bipartite graphs with arbitrary high girth.

5.1 On general case

Construction 2. *Let (C, U) be an instance of* SET COVER. *We construct a graph G as follows:*

- *For each set S_i, create a 3-path $sp_i = (v_i^0, v_i^1, v_i^2)$.*
- *For each element E_j, create a 4-path $ep_j = (u_j^0, u_j^1, u_j^2, u_j^3)$. We denote the edge $u_j^2 u_j^3$ as e_j.*
- *For each set S_i and each element $E_j \in S_i$, introduce a 3-path cp_j^i between v_i^1 and u_j^2) and a 2-path lp_j^i between v_i^2 and u_j^3.*
- *For each pair of elements E_j and $E_{j'}$, introduce the edge $t_{j,j'} = u_j^1 u_{j'}^1$.*

The paths ep_j, sp_i, cp_j^i and lp_k^i are called *element paths*, *set paths*, *cut paths* and *long paths*, respectively. An example of a graph produced by Construction 2 is depicted in Figure 3.

Clearly, the construction can be carried in polynomial time. In order to show that Construction 2 constitutes a reduction, we introduce the following lemmas.

Lemma 2. *Let (C, U) be an instance of* Set Cover *(resp.* $(1, k')$-Set Cover*) and let G be its graph resulting from Construction 2. The set $D_1(G)$ covers (resp. strongly covers) every edge of G' except the edges in $\{e_j \mid E_j \in U\}$.*

Proof. We start by describing the geodesics between each pair of vertices of $D_1(G)$ and how we can make the assignation them when there is some choice.

(a) For any pair of elements E_j and $E_{j'}$ of U, the distance between u_j^0 and $u_{j'}^0$ is 3 and the unique geodesic between these two vertices is described by $ep_j \sim t_{j,j'} \sim ep_{j'}$.
(b) For any set S_i of C and any element $E_j \in S_i$, the distance between u_j^0 and v_i^0 is 6 and the unique geodesic between these two vertices is described by $ep_j \sim cp_j^i \sim sp_i$.
(c) For any set S_i of C and any element $E_j \notin S_i$, the distance between u_j^0 and v_i^0 is 7 and the geodesics are described by $ep_j \sim t_{j,j'} \sim ep_{j'} \sim cp_{j'}^i \sim sp_i$, for any element $E_j \in S_i$. Since these geodesics contain only edges that are covered by the cases (a) or (b), we can assign any geodesic for this case.
(d) For any pair of sets S_i and $S_{i'}$ of C such that $S_i \cap S_{i'} \neq \emptyset$, the distance between v_i^0 and $v_{i'}^0$ is 8 and the geodesics between these two vertices are described by $sp_i \sim cp_j^i \sim cp_j^{i'} \sim sp_{i'}$ and $sp_i \sim lp_j^i \sim lp_j^{i'} \sim sp_{i'}$, for each $E_j \in S_i \cap S_{i'}$. Since, the edges of cp_j^i and $cp_j^{i'}$ are already covers by the case (b), it is better to assign a geodesic $sp_i \sim lp_j^i \sim lp_j^{i'} \sim sp_{i'}$, note that if $|S_i \cap S_{i'}| = 1$, then there only one such geodesic.
(e) For any pair of sets S_i and $S_{i'}$ of C such that $S_i \cap S_{i'} = \emptyset$, the distance between v_i^0 and $v_{i'}^0$ is 11 and the geodesics between these two vertices are described by $sp_i \sim cp_j^i \sim ep_j \sim t_{j,j'} \sim ep_{j'} \sim cp_{j'}^{i'} \sim sp_{i'}$, for any $E_j \in S_i$ and $E_{j'} \in S_{i'}$.

Remark that the e_j edges are not covered by any geodesic between the vertices of $D_1(G)$. We conclude by showing that every other edges are covered by $D_1(G)$, even if we fix a unique geodesic in the case where the intersection between two sets is at most one.

Let e be an edge in $E(G) \setminus \{e_j \mid E_j U\}$. If e belongs to a set path sp_i or a long path lp_i, then e is covered by the geodesic between v_i^0 and v_j^0. If e belongs to an element path ep_j or a cut path cp_j^i, then e is covered by the geodesic between v_i^0 and $u_{i,j}^0$. Finally, if e is a $t_{j,j'}$ edge, then e is covered by the geodesic between u_j^0 and $u_{j'}^0$. Hence, every edge of G is covered except the e_j edges.

In the following, let $Y_i^S \subset V(G)$ denote the set containing $sp_i \setminus \{v_i^0\}$ and every long path lp_j^i and cut path cp_j^i incident to cp_i minus vertices of every element path ep_j. Formally, $Y_i = (sp_i \setminus \{v_i^0\}) \cup \{(cp_j^i \cup lp_j^i) \setminus ep_j \mid \forall E_j \in S_i\}$. For each element $E_j \in U$, we also denote $Y_j^E = \{Y_i \mid E_j \in S_i\} \cup ep_j \setminus \{u_j^0\}$.

Lemma 3. *Let* (C, U) *be an instance of* Set Cover *and let* G *be its graph resulting from Construction 2. For each element* E_j*, every geodesic containing the edge* e_j *has an extremity in* Y_j^E.

Proof. Toward a contradiction, we suppose that there exists an edge e_j such that there is a geodesic g with extremities $x \notin Y_j^E$ and $y \notin Y_j^E$ that contains e_j. For simplicity, we denote the subpath $g[x, u_j^3]$ as g_x (we suppose that $e_j \notin g_x$). By hypothesis, g_x contains a long path lp_j^i such that $E_j \in S_i$. Since g_x can not have an extremity in Y_i^S, g_x contains a vertex x' in some element path ep_k, such that $E_k \in S_i$. The subpath $g[x', u_j^2]$ can be described either by $cp_k^i \sim sp_i \sim cp_j^i \sim ep_j$ or by $lp_k^i \sim sp_i \sim lp_j^i \sim ep_j$. Let g' be the path between x' and u_j^2 described by $ep_k \sim t_{k,j} \sim ep_j$. By construction, we have $|g'| < |g[x', u_j^2]|$. Thus replacing $g[x', u_j^2]$ by g' in g constructs a path between x and y that is shorter than g, contradicting that g is a geodesic. Hence, X contains at least one vertex in Y_j^E.

In order to easily produce a set cover in G from a (strong) edge geodetic set X of G', we need X to respect a certain property. Hence, we use the following lemma.

Lemma 4. *Let* (C, U) *be an instance of* Set Cover *(resp.* $(1, k')$-Set Cover*) and* G *its graph resulting from Construction 2. Let* $X \subseteq V(G)$ *be an edge geodetic set (resp. strong edge geodetic set) of* G*. It is possible to construct an edge geodetic set (resp. a strong edge geodetic set)* X' *of* G *such that* $|X'| \leq |X|$ *and*

$$X' \subseteq \{v_i^2 \mid S_i \in C\} \cup D_1(G)$$

.

Proof. First by Property 1, X contains any vertices of $D_1(G)$. Therefore, by Lemma 2, every edge is (strongly) covered by $D_1(G)$ except the e_j edges. Note that selecting a vertex v_i^2 ensures to cover every e_j edges (with $E_j \in S_i$) since the unique geodesic between v_i^1 and u_j^0 is described by $lp_j^i \sim ep_j$.

We show how to transform X so that it respects lemma's property. For each edge e_j, by Lemma 3, there is a vertex x in the intersection $X \cap Y_j^E$. If $x \in Y_i^S$ for some S_i such that $E_j \in S_i$, we replace x in X by v_i^2. Since x belongs to any set Y_k^E such that $S_k \in E_i$, every edge previously covered by x is still covered. If $x \in ep_j$ then it is only used to cover e_j. Thus, we can replace x arbitrarily by any vertex v_i^2 such that $E_j \in S_i$. Finally, there exists a vertex that does not belong to $\{v_i^2 \mid S_i \in C\} \cup D_1(G)$, we can remove it from X. Hence, we obtain a solution as thought.

Lemma 5. *Let* (C, U) *be an instance of* Set Cover *(resp.* $(1, k')$-Set Cover*) and* G *its graph resulting from Construction 2. Then the instance* (C, U) *contains a set cover of size* k *if and only if* G *contains an edge geodetic set (resp. strong edge geodetic set) of size* $k + |C| + |U|$.

Proof.
- Let $C' \subseteq C$ be a set cover of size k of (C, U) and consider the (strong) edge geodetic set $X = D_1(G) \cup \{v_i^2 \mid S_i \in C'\}$. By Lemma 2, every edges

in $E(G) \setminus \{e_j \mid E_j \in U\}$ are (strongly) covered. Let E_j be an element of U, there is a set $S_i \in S'$ such that $E_j \in S_i$. Thus, $v_i^2 \in X$ and therefore the edge e_j is covered by the unique geodesic between v_i^2 and $u_{i,j}^0$. Since every element of U appears in S', every e_j edges of G are also (strongly) covered. Hence, we produce a (strong) edge geodetic set of size $k + |C| + |U|$.
– Let $X \subset V(G)$ be a (strong) edge geodetic set of size $k+|C|+|U|$ of G that respects the property of Lemma 4. Consider the set cover $C' = \{S_i \mid v_i^2 \in X\}$. Let E_j be an element of U and suppose that it does not belong to a set of C'. In that case, no vertex v_i^2 such that $E_j \in S_i$ belongs to X. By Lemma 3, the edge e_j is not covered by X, contradicting X being a (strong) edge geodetic set. Hence, every element of U is contained in a set of C' and, we construct a set cover of size k.

5.2 On subcubic bipartite graphs

We now extend the previous result to subcubic and bipartite graphs. First, we show that the result holds in graph with maximum degree three. We introduce the following construction.

Construction 3. *Given a graph G, a vertex $u \in V(G)$, a set of non-adjacent neighbours $N^0 = \{v_0^0, \dots, v_{k-1}^0\} \subseteq N(u)$ and an integer $h > \log k$, emplace a h-pyramid $Py(h, u, N^0)$ consists of removing all edges between u and N^0 and replacing them with the following subgraph. For each $0 < i < h$, construct recursively the sets N^i:*
– *create $t = \lceil |N^{i-1}|/2 \rceil$ vertices $v_0^i, \dots, v_t^i$, and*
– *introduce the edges $v_{t'}^i v_{2t'}^{i-1}$, and $v_{t'}^i v_{2t'+1}^{i-1}$ (if $v_{2t'+1}^{i-1}$ exists) for each $t' < t$.*
Finally, introduce the edge uv_0^{h-1} (N^{h-1} consists of a single vertex since $h > \log k$).

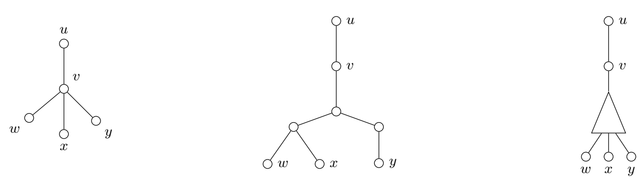

Fig. 4. Example of a 3-pyramid $Py(3, v, \{w, x, y\})$ produced by Construction 3. **Left:** v and its neighbours in the original graph. **Center:** emplaced 3-pyramid. **Right:** Representation of the pyramid used in Figure 3.

Let $Py(h, v, N)$ be a h-pyramid. We can make the following observations.
– The maximum degree of $Py_h(h, v, N)$ is three.
– Let $n_1, n_2 \in N$, the distance between n_1 and n_2 in $Py_h(h, v, N)$ is between 2 and $2h$ and the distance between v and n_1 or n_2 is h.

We now use the previous structure to modify Construction 2 as follows.

Construction 4. *Let (C, U) be an instance of* Set Cover *and G be its graph resulting from Construction 2. Let $h > \log \Delta(G)$ be an integer. We modify G as follows:*

- *for each set S_i and each integer $k \in \{1, 2\}$, emplace a h-pyramid $Py(h, v_i^k, N(v_i^k) \setminus \{v_i^{k-1}\})$,*
- *for each element E_j and each integer $k \in \{1, 2, 3\}$, emplace a h-pyramid $Py(h, u_j^k, N(u_j^k) \setminus \{u_j^{k-1}\})$, and*
- *replace each edge of G that does not belong to a h-pyramid by a path of length h.*

Note that the resulting graph has maximum degree three. Moreover, if k is odd then the resulting graph is bipartite. Finally, by taking an arbitrary high value of k, the resulting graph has an arbitrary high girth. We use a similar vocabulary than for Construction 2: an *element tree et_j* is the tree induced by the vertices of ep_j in the h-pyramids emplaced in it in the original graph. A *set tree st_i* is defined the same way. For each element E_j, the h-path that replaces the edge e_j is denoted p_j. An example of a graph produced by Construction 4 is depicted in Figure 3. Since Construction 4 multiplies the length of every path of Construction 2 by h, we can adapt Lemmas 2 to 4 to it by replacing ep_j by et_j and sp_i by st_i in the geodesics descriptions.

Lemma 6. *Let (C, U) be an instance of* Set Cover *(resp.* $(1, k')$-Set Cover*) be a connected graph and G its graph resulting from Construction 4. The set $D_1(G)$ covers (resp. strongly cover) every edge and vertex of G except edges and vertices in $\{p_j \mid E_j \in U\}$.*

Proof. By construction, geodesics between vertices of $D_1(G)$ can be described the same way as for Lemma 2. The result follows.

In the following Y_i^S and Y_j^E are defined in the same way as for Construction 2 (by taking st_i instead of sp_i and et_j instead of ep_j). Notice that the geodesic between a vertex v^1 and a vertex u_j^0 is described by $st_i \sim lp_j^i \sim et_j$ and contains the path p_j. Hence, using the same arguments as in Lemmas 3 and 4, we can show the two following results.

Lemma 7. *Let (C, U) be an instance of* Set Cover *and G be its graph resulting from Construction 4. For each element E_j, every geodesic containing a vertex of p_j has an extremity in Y_j^E.*

Lemma 8. *Let (C, U) be an instance of* Set Cover *(resp.* $(1, k')$-Set Cover*) and G be its graph resulting from Construction 4. Let $X \subseteq V(G)$ be a geodetic or an edge geodetic set of G' (resp. strong geodetic or a strong edge geodetic set). It is possible to construct a geodetic or an edge geodetic set (resp. strong geodetic or a strong edge geodetic set) X' of G such that $|X'| \leq |X|$ and*

$$X' \subseteq \{v_i^2 \mid S_i \in C\} \cup D_1(G).$$

Using the same idea as for Lemma 5, we can now show that Construction 4 constitutes a reduction: if a path p_j is (strongly) covered, then there is a vertex v_i^2, such that $E_j \in S_i$, is selected. Thus, given a solution for a geodetic problem X the set $\{S_i \mid v_i^2 \in X\}$ is a set cover of G. Hence, we obtain the following result.

Lemma 9. *Let (C, U) be an instance of* Set Cover *(resp.* $(1, k')$-Set Cover*) and G its graph resulting from Construction 2. Then the instance (C, U) contains a set cover of size k if and only if G contains an edge geodetic set and a geodetic set (resp. strong edge geodetic set and a strong geodetic set) of size $k + |C| + |U|$.*

6 Non-approximability

In this section, we use the results of the previous section to find hardness of approximation results for the geodetic problems.

6.1 Strong geodetic set and strong edge geodetic set

First, recall the definition of *L-reduction* between two hard problems Π and Π' (with respective cost functions val_Π and $val_{\Pi'}$), as described by Papadimitriou and Yannakakis [19]. Let $OPT_\Pi(x)$ and $OPT_{\Pi'}(x)$ be the optimal value of val_Π and $val_{\Pi'}$ on an instance x, respectively. An L-reduction consists of polynomial-time computable functions f and g such that, for each instance x of Π, $f(x)$ is an instance of Π' and for each feasible solution y' for $f(x)$, $g(y')$ is a feasible solution for x. Moreover, there are constants $\alpha_1, \alpha_2 > 0$ such that:

1. $OPT_{\Pi'}(f(x)) \leq \alpha_1 \cdot OPT_\Pi(x)$ and
2. $|val_\Pi(g(y')) - OPT_\Pi(x)| \leq \alpha_2 \cdot |val_{\Pi'}(y') - OPT_{\Pi'}(f(x))|$.

Using Construction 4, we obtain an L-reduction with $\alpha_1 = (2k' + 2)$ and $\alpha_2 = 1$.

Lemma 10. *Let $\rho_{k'}$ be the best possible polynomial time approximation factor of* $(1, k')$-Minimum Set Cover*. Then* Strong Geodetic Number *and* Strong Edge Geodetic Number *cannot be approximated with a factor better than*

$$1 + \frac{\rho_{k'} - 1}{2k' + 2},$$

in subcubic bipartite graphs with arbitrary high girth.

Proof. Let (C, U) be an instance of $(1, k')$-Minimum Set Cover and G its graph resulting from Construction 4. Suppose there is a polynomial-time approximation algorithm for Strong Geodetic Number (resp. Strong Edge Geodetic Number) in subcubic bipartite graphs with arbitrary high girth and let X_{app} be a strong geodetic set (resp. strong edge geodetic set) computed by this algorithm in G. We suppose that X_{app} respects the property of Lemma 8. Let $C'_{app} = \{S_i \mid v_i^2 \in X_{app}\}$. Using the same argument as in Lemma 9, we can show that C'_{app} is a set cover of (C, U) and $|X_{app}| = |C| + |U| + |C'_{app}|$. Let $C'_{opt} \subset C$ be a set cover of (C, U) and let X_{opt} be a minimum strong geodetic set (resp. strong edge geodetic set) in G. Similarly, we can show that $|X_{opt}| = |C| + |U| + |C'_{opt}|$. Hence, we have

$$|X_{app}| - |X_{opt}| = |C'_{app}| - |C'_{opt}|. \tag{1}$$

Moreover, since the intersection between two sets contains at least two elements and the size of each set is bounded by k', we deduce

$$|U| \leq k' \cdot |C'_{opt}| \text{ and } |C| \leq (k'+1) \cdot |C'_{opt}|$$

which leads to

$$|X_{opt}| \leq (2k'+2) \cdot |C'_{opt}|. \tag{2}$$

Thus we construct a L-reduction with $\alpha_1 = (2k'+2)$ and $\alpha_2 = 1$. We conclude,

$$\begin{aligned}|X_{app}| &\overset{(1)}{=} |C'_{app}| - |C'_{opt}| + |X_{opt}| \\ &\geq (\rho_{k'} - 1) \cdot |C'_{opt}| + |X_{opt}| \\ &\overset{(2)}{\geq} (1 + \frac{\rho_{k'} - 1}{2k'+2}) \cdot |X_{opt}|.\end{aligned}$$

Since Minimum Vertex Cover with bounded maximum degree k' is a particular case of $(1, k')$-Minimum Set Cover, we can pick the value of k' (and so the corresponding best-known value $\rho_{k'}$) that maximize the previous inapproximation ratio. Thus, since Berman and Karpinski showed that Minimum Vertex Cover cannot be approximated with a factor better than $^{79}/_{78}$ in graphs with maximum degree four [4], we obtain the following result.

Corollary 2. Strong Geodetic Number *and* Strong Edge Geodetic Number *cannot be approximated with a factor better than* $^{781}/_{780}$ *in subcubic bipartite graphs with arbitrary high girth.*

6.2 Geodetic set and edge geodetic set

Now, we provide approximation lower bounds for Geodetic Number and Strong Geodetic Number. We apply the following modification to Construction 4.

Construction 5. *Let (C, U) be an instance of* Set Cover*, G be its graph produced by Construction 4 and $k > |V(G)|$ be an integer. We construct a graph G' as follows:*

- *create k disjoint copies $\{G_1, \ldots, G_k\}$ of G,*
- *for each vertex x of $D_1(G)$,*
  - *create an edge $s_x^0 s_x^1$,*
  - *for each $G_\ell \in \{G_1, \ldots, G_k\}$ and for each vertex $x \in D_1(G_\ell)$, construct a k-path p_x^ℓ between x and s_x^1, and*
  - *emplace a k-pyramid $Py(k, s_x^1, N(s_x^1) \setminus \{s_x^0\})$.*

Notice that the resulting graph has maximum degree three. For simplicity, we denote the k-pyramide $Py(k, s_x^1, N(s_x^1) \setminus \{s_x^0\})$ as $Py[x]$.

Lemma 11. *Let (C, U) be an instance of* Set Cover *and let G' be its graphs resulting from Construction 5. The set $D_1(G')$ covers every edge and vertex of G' except edges and vertices in $\{p_j \mid E_j \in U, G_\ell \in \{G_1, \ldots, G_k\}, p_j \in G_\ell\}$.*

Proof. Let G be the graph of (C, U) produced Construction 4. Let $x \in D_1(G)$ and G_ℓ in $\{G_1, \ldots, G_k\}$. The shortest path between G_ℓ and s_x^0 has length $2k$ and walks through p_k^ℓ and $Py[x]$. Let x_1 and x_2 be two vertices of $D_1(G)$ and x_1' and x_2' be their corresponding vertices in G_ℓ. Let g_ℓ be a geodesic between x_1' and x_2', note that g_ℓ is entirely contained in G_ℓ since every path leaving G_ℓ has length $k > |V(G_\ell)|$. Suppose that $g = s_{x_1}^0 s_{x_1}^1 \sim Py[x_1] \sim p_{x_1}^\ell \sim g_\ell \sim p_{x_2}^\ell \sim Py[x_2] \sim s_{x_2}^0 s_{x_2}^1$ is not a geodesic between $s_{x_1}^0$ and $s_{x_2}^0$. Then, let g' be a geodesic between $s_{x_1}^0$ and $s_{x_2}^0$. If g' contains two subgraph G_ℓ and $G_{\ell'}$, then since the distance between G_ℓ and $G_{\ell'}$ is $2k$, the length of g' is at least $6k$ and then g is shorter than g'. Thus, g' contains only one subgraph G_ℓ and then, it can be described by $s_{x_1}^0 s_{x_1}^1 \sim Py[x_1] \sim p_{x_2}^\ell \sim g_\ell' \sim p_{x_2}^\ell \sim Py[x_2] \sim s_{x_2}^0 s_{x_2}^1$, where g_ℓ' is a path between x_1' and x_2' in G_ℓ, but then it contradicts g_ℓ being a geodesic. Hence, g is a geodesic between $s_{x_1}^0$ and $s_{x_2}^0$. From that, we can conclude that $D_1(G')$ covers exactly the same edges and vertices than $D_1(G)$ in each G_ℓ. Hence, by Lemma 6, every edge and vertex in any G_ℓ is covered, except edges and vertices in $\{p_j \mid E_j \in U\}$.

It remains to show that edges and vertices in every $Py[x]$ and p_x^ℓ are also covered. First, since for every $x \in D_1(G)$ and every G_ℓ, there is a geodesic g with extremity s_x^0 that walks through G_ℓ, then g contains p_x^ℓ and thus every edge and vertex of p_x^ℓ is covered. Moreover, since every p_x^ℓ is adjacent to a distinct leaf of $Py[x]$, the set of geodesic with extremity s_x^0 covers every edge and vertex of $Py[x]$. The result follows.

Lemma 12. *Let (C, U) be an instance of* Set Cover *and let G' be its graph resulting from Construction 5. Let $X \subseteq V(G')$ be a geodetic set or an edge geodetic set of G'. It is possible to construct a geodetic set or an edge geodetic set X' of G' such that $|X'| \leq |X|$ and*

$$X' \subseteq \{v_i^2 \mid S_i \in C, G_\ell \in \{G_1, \ldots, G_k\}, v_i^2 \in V(G_\ell)\} \cup D_1(G').$$

Proof. Since for each $G_\ell \in \{G_1, \ldots, G_k\}$, the result of Lemma 7 holds in G_ℓ, we can use the same technique as for Lemma 8 to obtain a set X' such that $X' \cap V(G_\ell) = \{v_i^2 \mid S_i \in C\}$. Then, the result follows.

Set Cover is hard to approximate with a factor better than a logarithmic function [9]. Therefore, we can transfer the lower bounds of approximation of Set Cover to Geodetic Number and Edge Geodetic Number. This result is in addition to the one proved by Chakraborty *et al.* [7].

Theorem 3. Geodetic Number *and* Edge Geodetic Number *are LOG-$\mathcal{APX}$-hard, even in bipartite subcubic graphs with arbitrary high girth.*

Proof. Let (C, U) be an instance Set Cover graph and G' be its graph produced by Construction 5. Suppose that we have a polynomial-time approximation algorithm A to compute a geodetic set or an edge geodetic set of a graph. We denote by X the geodetic set obtained by A in G' and we suppose that X respects Lemma 12 property. We can suppose that $|X \cap V(G_\ell)|$ has the same value for any $G_\ell \in \{G_1, \ldots, G_k\}$, since otherwise it suffixes to transpose the solution with the minimum $|X \cap V(G_\ell)|$ value to the every other G_ℓ. Let $C' = \{S_i \mid v_i^2 \in X \cap V(G_1)\}$. Using the same argument as in the proof of Lemma 5, we can show that C is a set cover of (C, U), and that $|X| = k|C'| + |C| + |U|$.

Let X_{opt} be a minimum geodetic set or a minimum edge geodetic set of G' and C'_{opt} be a minimum set cover of (C, U). Similarly, we can show that $|X_{opt}| = k|C'_{opt}| + |U| + |C|$. We have:

$$\frac{|X|}{|X_{opt}|} = \frac{k|C'| + |U| + |C|}{k|C'_{opt}| + |U| + |C|} \tag{3}$$

As k tends to $+\infty$ we deduce that $\frac{|X|}{|X_{opt}|}$ converges to $|C'|/|C'_{opt}|$. Since Set Cover cannot be approximated with a factor better than a logarithmic function, we deduce that A cannot have an approximation factor better than a logarithmic function.

7 Strong Geodetic Number on complete multipartite graphs

First, remark that geodesics of complete multipartite graphs are easy to determine: for any pair of vertices which are not in the same part, the edge between them is the unique shortest path between them. For a pair of vertices which are in the same part, the shortest paths between them are all the paths of length two between them with all the vertices not in this part as middle vertices.

In this section, we develop a polynomial algorithm which computes the strong geodetic number of a complete multipartite graph. The algorithm is based on dynamic programming where we not only look after a minimum strong geodetic set of vertices covering all the graph, but we look after all sets of vertices maximizing the number of pairs not used to cover other vertices among sets with some fixed parameters.

Let $K_{n_1,\ldots,n_r}$ denotes a complete multipartite graph whose parts are noted $X_1, \ldots, X_r$ such that $|X_i| = n_i$ for every $i \in \{1, \ldots, r\}$. We denote $N_i = \sum_{j=1}^{i} n_j$ for every $i \in \{1, \ldots, r\}$ and $K_{n_1,\ldots,n_i}$ by G_i.

Definition 1. *A* selection *of $K_{n_1,\ldots,n_r}$ is a set of selected vertices S in which we pick a set of pairs of non-adjacent vertices C to cover some non-selected vertices. Formally, a selection is a triplet (S, C, f), where*

- $S \subseteq V$,
- $C \subseteq \bigcup_{j=1}^{r} \mathcal{P}_2(S \cap X_j)$ *and,*

– *$f : C \to V \setminus S$ is an injective map such that $\forall c \in C \bigcap \mathcal{P}_2(S \bigcap X_i), f(c) \notin X_i$ (i.e. two vertices of X_i can not cover another vertex of X_i).*

Given a selection $s(S, C, f)$, we denote by

– $s(S, C, f) = |S|$, the number of selected vertices,
– $r(S, C, f) = |V \setminus (S \sqcup f(C))| = n - s(S, C, f) - |C|$, the number of vertices that are neither selected nor covered, and
– $d(S, C, f) = |\cup_{j=1}^{n} \mathcal{P}_2(S \cap S_j) \setminus C| = \sum_{j=1}^{n} \binom{S \cap X_j}{2} - |C|$, the number of pairs of non-adjacent vertices that are not in C.

We say that a selection (S, C, f) is nice if it maximizes the number $d(S, C, f)$ among all selection with the same size $s(S, C, f)$ and the same number $r(S, C, f)$. In our dynamic programming approach, we construct every nice selections in every subgraph G_i, successively. To construct a selection in G_i from a selection of G_{i-1}, we use the following lemma.

Lemma 13. *Let (S', C', f') a selection of G_{i-1}. Let three integers k, u and q such that $0 \leq k \leq n_i$, $0 \leq u \leq \min(n_i - k, d(S', C', f'))$ and $0 \leq q \leq \min(\binom{k}{2}, r(S', C', f'))$. There exists a selection (S, C, f) of G_i such that*

$$\begin{cases} s(S, C, f) & = s(S', C', f') + k \\ r(S, C, f) & = r(S', C', f') + n_i - k - q - u \\ d(S, C, f) & = d(S', C', f') + \binom{k}{2} - q - u. \end{cases}$$

In other words, we select k vertices in X_i, cover u vertices in X_i and cover q vertices in G_{i-1}.

Proof. Let T be any subset of X_i of size k. Let Q be any subset of $\mathcal{P}_2(T)$ of size q. Let U be any subset of $\bigcup_{j=1}^{i-1} \mathcal{P}_2(S' \cap X_j) \setminus C'$ of size u. Let Q' be any subset of $\bigcup_{j=1}^{i-1} X_j \setminus (S' \cup f'(C'))$ of size q. Let U' be any subset of $X_i \setminus T$ of size u.

We define

$$\begin{aligned} S &= S' \cup T \\ C &= C' \cup Q \cup U \\ f &: C \to V \setminus S \end{aligned}$$

where the function f is defined as follows. On C', $f_{|C'} = f'$. As U (resp. Q) and U' (resp. Q') are of the same size, we can define $f_{|U}$ (resp. $f_{|Q}$) as any bijection between U and U' (resp. Q and Q').

Let us show that (S, C, f) is a selection of G_i. As $S' \subseteq \bigcup_{j=1}^{i-1} X_j$ and $T \subseteq X_i$, we deduce that $S \subseteq \bigcup_{j=1}^{i} X_j$. As $C' \subseteq \bigcup_{j=1}^{i-1} \mathcal{P}_2(S \cap X_j)$, $Q \subseteq \mathcal{P}_2(S \bigcap X_i)$ and $U \subseteq \bigcup_{j=1}^{i-1} \mathcal{P}_2(S \cap X_j)$, we deduce that $C \subseteq \bigcup_{j=1}^{i} \mathcal{P}_2(S \cap X_j)$.

By definition of f, $f_{|C'}$, $f_{|Q}$ and $f_{|U}$ are injective. Let $x, x' \in C$ such that $f(x) = f(x')$. As $f(U) \subseteq X_i$ and $f(Q)$ and $f(C') \subseteq \bigcup_{j=1}^{i-1} X_j$, we deduce that we can suppose that $x \in Q$ and $x' \in C'$. As $f(x) \in Q'$, we have $f(x) \notin f'(C')$. It contradicts the fact that $f(x) = f(x') \in f'(C')$ because $x' \in C'$. We conclude that $x = x'$ and that f is injective.

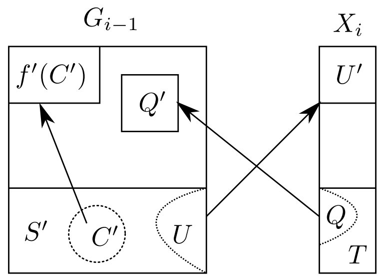

Fig. 5. Description of the creation of selection (S, C, f) of G_i from the selection (S', C', f') of G_{i-1}. An arrow between two set indicates that pairs of vertices of the source set are used to cover vertices of the target set.

Let $a \in C \cap \mathcal{P}_2(S \cap X_i) = Q$. Then $f(a) \in Q' \subseteq \bigcup_{j=1}^{i-1} X_j$. Thus, $f(a) \notin X_i$. Let $a \in C \cap \mathcal{P}_2(S \cap X_j) = C' \cup U$ for any $j < i$. If $a \in C'$, then $f(a) = f'(a) \notin X_j$ because (S', C', f') is a selection of G_{i-1}. Otherwise, $a \in U$ and $f(a) \in X_i$ and thus $f(a) \notin X_j$.

We conclude that (S, C, f) is a selection of G_i.

As $S = S' \sqcup T$, we deduce that $s(S, C, f) = s(S', C', f') + |T| = s(S', C', f') + k$. As $C = C' \sqcup Q \sqcup U$, we deduce that $r(S, C, f) = N_i - |S| - |C| = n_i + N_{i-1} - |S'| - k - |C'| - |Q| - |U| = n_i - k - q - u + r(S', C', f')$. Furthermore

$$\begin{aligned} d(S, C, f) &= \sum_{j=1}^{i} \binom{S \cap X_j}{2} - |C| \\ &= \sum_{j=1}^{i-1} \binom{S' \cap X_j}{2} + \binom{|T|}{2} - |C'| - q - u \\ &= d(S', C', f') + \binom{k}{2} - q - u \end{aligned}$$

Since the number of selections can be exponential, we only keep nice selections. In order to do that, we adapt the result of Lemma 13 as follows.

Lemma 14. *Let (S, C, f) be a nice selection of G_i, then there exists a nice selection (S', C', f') of G_{i-1} and numbers k, u, q such that $0 \leq k \leq n_i$, $0 \leq q \leq \min(\binom{k}{2}, r(S', C', f'))$ and $0 \leq u \leq \min(n_i - k, d(S', C', f'))$ and*

$$\begin{cases} s(S, C, f) = s(S', C', f') + k \\ r(S, C, f) = r(S', C', f') + n_i - u - q - k \\ d(S, C, f) = d'(S', C', f') + \binom{k}{2} - q - u \end{cases}$$

Proof. We define $S' = S \bigcap G_{i-1}$, $C' = C \bigcap G_{i-1} \setminus f^{-1}(X_i)$ and $f' = f_{|C'}$. Let us prove that (S', C', f') is a nice selection of G_{i-1} satisfying the announced equalities.

Let us show that $\forall c \in C' \bigcap \mathcal{P}_2(S \bigcap X_i), f'(c) \notin X_i$. Let $c \in C' \bigcap \mathcal{P}_2(S \bigcap X_i)$, suppose by contradiction that $f'(c) \in X_i$. As $f'(c) = f(c)$, we deduce that $c \in f^{-1}(X_i)$. This contradicts the definition of C'. Thus, $f'(c) \notin X_i$. As f is injective, f' is also injective and, we deduce that (S', C', f') is a selection of G_{i-1}.

We define $k = |S \cap X_i|$, $q = |C \cap \mathcal{P}_2(S \cap X_i)|$ and $u = |f^{-1}(X_i)|$. As $S = S' \sqcup (S \cap X_i)$, we have $s(S) = s(S') + k$. As we have $C = C' \sqcup f^{-1}(X_i) \sqcup (C \cap \mathcal{P}_2(S \cap X_i)$, thus $|C| = |C'| + u + q$. We deduce that $r(S) = N_i - |S| - |C| = n_i + N_{i-1} - s(S', C', f') - k - |C'| - u - q = r(S', C', f') + n_i - k - u - q$. Furthermore,

$$\begin{aligned}
d(S, C, f) &= \sum_{j=1}^{i} \binom{S \cap X_j}{2} - |C| \\
&= \sum_{j=1}^{i-1} \binom{S' \cap X_j}{2} + \binom{k}{2} - |C'| - q - u \\
&= d(S', C', f') + \binom{k}{2} - q - u
\end{aligned}$$

We have $k = |S \cap X_i| \leq |X_i| \leq n_i$. We have $q \leq |\mathcal{P}_2(S \cap X_i)| \leq \binom{k}{2}$. As $r(S, C, f) = r(S', C', f') + n_i - k - u - q$, we have $q = r'(S', C', f') - r(S, C, f) + n_i - k - u \leq r'(S', C', f')$ as we have $n_i \leq r(S, C, f) + k + u$ as $r(S, C, f)$ is the size of the set $\bigcup_{j=1}^{i} X_j \setminus (S \sqcup f(C))$ and as $X_i \setminus ((S \cap X_i) \sqcup f(f^{-1}(X_i))) \subseteq \bigcup_{j=1}^{i} X_j \setminus (S \sqcup f(C))$. Thus, $q \leq r(S', C', f')$.

As $(S \cap X_i) \sqcup f(f^{-1}(X_i)) \subseteq X_i$, we deduce that $k + u \leq n_i$ (as f is injective), thus $u \leq n_i - k$. As $d(S, C, f)$ is the size of the set $\bigcup_{j=1}^{i} \mathcal{P}_2(S \cap X_j) \setminus C$, we have $\mathcal{P}_2(S \cap X_i) \setminus (C \cap \mathcal{P}_2(S \cap X_i)) \subseteq \bigcup_{j=1}^{i} \mathcal{P}_2(S \cap X_j) \setminus C$. Thus, $\binom{k}{2} - q \leq d(S, C, f)$ and so, $u = d(S', C', f') - d(S, C, f) + \binom{k}{2} - q \leq d(S', C', f')$.

We conclude that k, q, u satisfy the following inequalities:

$$\begin{aligned}
&0 \leq k \leq n_i \\
&0 \leq q \leq \min\left(\binom{k}{2}, r(S', C', f')\right) \\
&0 \leq u \leq \min(n_i - k, d(S', C', f'))
\end{aligned}$$

Let us show that (S', C', f') is a nice selection of G_{i-1}. Suppose by contradiction that there exists a selection (S'', C'', f'') of G_{i-1} such that $s(S'', C'', f'') = s(S', C', f')$ and $r(S'', C'', f'') = r(S', C', f')$ and that $d(S'', C'', f'') > d(S', C', f')$. According to Lemma 13, there exists a selection (Sp, Cp, fp) of G_i such that

$$\begin{cases}
s(Sp, Cp, fp) & = s(S'', C'', f'') + k \\
r(Sp, Cp, fp) & = r(S'', C'', f'') + n_i - k - q - u \\
d(Sp, Cp, fp) & = d(S'', C'', f'') + \binom{k}{2} - q - u
\end{cases}$$

As $s(S'', C'', f'') = s(S', C', f')$, $r(S'', C'', f'') = r(S', C', f')$ and $d(S'', C'', f'') > d(S', C', f')$, we obtain

$$s(Sp, Cp, fp) = s(S, C, f)$$
$$r(Sp, Cp, fp) = r(S, C, f)$$
$$d(Sp, Cp, fp) > d(S, C, f)$$

Which contradicts that (S, C, f) is a nice selection of G_i. We conclude that (S', C', f') is a nice selection of G_{i-1}.

We denote by $d(i, j, r)$ the maximum of $d(S, C, f)$ for any selection (S, C, f) of G_i such that $s(S, C, f) = j$ and $r(S, C, f) = r$. This quantity is set to $-\infty$ if no such selection of G_i exists.

Lemma 15. *For any integers i, s, r and integers k, u, q we define the following quantities:*

$$s' = s - k, \quad r' = r - n_i + u + q + k \quad \text{and} \quad d' = d(i-1, s', r')$$

We deduce that: $d(i, s, r) = \max\left\{ (d') + \binom{k}{2} - q - u \,\middle|\, \begin{smallmatrix} 0 \le k \le n_i \\ 0 \le u \le \min(n_i - k, d') \\ 0 \le q \le \min(\binom{k}{2}, r') \end{smallmatrix} \right\}$

Proof. According to Lemma 13, we have

$$d(i, s, r) \le \max\left\{ (d') + \binom{k}{2} - q - u \,\middle|\, \begin{matrix} 0 \le k \le n_i \\ 0 \le u \le \min(n_i - k, d') \\ 0 \le q \le \min(\binom{k}{2}, r') \end{matrix} \right\}$$

as for any selection (S', C', f') of G_{i-1} such that $s(S', C', f') = s'$, $r(S', C', f') = r'$ and $d(S', C', f') = d'$, we can create a selection (S, C, f) of G_i such that $s(S, C, f) = s$, $r(S, C, f) = r$ and $d(S, C, f) = d' + \binom{k}{2} - q - u$.

According to Lemma 14, we have

$$d(i, s, r) \ge \max\left\{ (d') + \binom{k}{2} - q - u \,\middle|\, \begin{matrix} 0 \le k \le n_i \\ 0 \le u \le \min(n_i - k, d') \\ 0 \le q \le \min(\binom{k}{2}, r') \end{matrix} \right\}$$

as for any selection (S, C, f) of G_i such that $s(S, C, f) = s$, $r(S, C, f) = r$ and $d(S, C, f) = d(i, s, r)$ we can create a nice selection (S', C', f') of G_{i-1} such that $s(S', C', f') = s'$, $r(S', C', f') = r'$ and $d(S', C', f') = d(S, C, f) - \binom{k}{2} + q + u$. We deduce the equality between the two quantities.

From previous lemma we deduce the following theorem.

Theorem 4. *There exists an algorithm in $O(n^8)$ computing the geodetic number of a complete multipartite graph with n vertices.*

Proof. The algorithm 2 consists mainly in six independent "for" loops. Four of them are of length at most n and two of them are of length at most n^2. The complexity of the algorithm is therefore in $O(n^8)$.

Notice that a complete multipartite graph can be described with the list of integers $n_1, \ldots, n_k$. In that case, the dynamic programming that we described is not polynomial if the values of the n_i are exponential. Thus, if we formulate Edge Geodetic Number on complete multipartite graph as a specific problem on this class, the question whether such a problem is weak $\mathcal{NP}$-hard or not is open.

8 Conclusion

In this paper, we investigated the hardness of the approximation of the geodetic set problems. Given our approximation lower bound for Geodetic Number and Edge Geodetic Number, the question of the existence of a $O(\log(n))$-approximation algorithm seems natural. We also proved that deciding whether a set admits a geodesic assignation NP-hard. Therefore, a second question arises: is it also hard to decide whether a set of vertices is a strong geodetic set. We also give a tight 2-approximation of this problem. Finding a lower bound for this problem is probably a good question for further work. Finally, for Strong Geodetic Number, we proved that it was polynomial on complete multipartite graphs. What about other graph classes?

References

1. Anand, B.S., Changat, M., Chandran, S.U.: The edge geodetic number of product graphs. In: Conference on Algorithms and Discrete Applied Mathematics. pp. 143–154. Springer (2018)
2. Arokiaraj, A., Klavzar, S., Manuel, P.D., Thomas, E., Xavier, A.: Strong geodetic problems in networks. Discuss. Math. Graph Theory **40**(1), 307–321 (2020)
3. Atici, M.: On the edge geodetic number of a graph. International Journal of Computer Mathematics **80**(7), 853–861 (2003)
4. Berman, P., Karpinski, M.: On some tighter inapproximability results (extended abstract). In: Automata, Languages and Programming, 26th International Colloquium, ICALP'99, Prague, Czech Republic, July 11-15, 1999, Proceedings. pp. 200–209 (1999)
5. Bresar, B., Kovse, M., Tepeh, A.: Geodetic sets in graphs. In: Dehmer, M. (ed.) Structural Analysis of Complex Networks, pp. 197–218. Birkhäuser / Springer (2011)
6. Chakraborty, D., Das, S., Foucaud, F., Gahlawat, H., Lajou, D., Roy, B.: Algorithms and Complexity for Geodetic Sets on Planar and Chordal Graphs. In: Cao, Y., Cheng, S.W., Li, M. (eds.) 31st International Symposium on Algorithms and Computation (ISAAC 2020). Leibniz International Proceedings in Informatics (LIPIcs), vol. 181, pp. 7:1–7:15. Schloss Dagstuhl–Leibniz-Zentrum für Informatik (2020)
7. Chakraborty, D., Foucaud, F., Gahlawat, H., Ghosh, S.K., Roy, B.: Hardness and approximation for the geodetic set problem in some graph classes. In: Changat, M., Das, S. (eds.) Algorithms and Discrete Applied Mathematics - 6th International Conference, CALDAM 2020, Hyderabad, India, February 13-15, 2020, Proceedings. Lecture Notes in Computer Science, vol. 12016, pp. 102–115. Springer (2020)
8. Darmann, A., Döcker, J.: On simplified NP-complete variants of monotone 3-sat. Discret. Appl. Math. **292**, 45–58 (2021)

9. Dinur, I., Steurer, D.: Analytical approach to parallel repetition. In: Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing. p. 624–633. STOC '14, Association for Computing Machinery, New York, NY, USA (2014)
10. Doughat, A., Kong, M.: Computing geodetic bases of chordal and split graphs. Journal Of Combinatorial Mathematics And Combinatorial Computing **22**, 67–78 (1996)
11. Douthat, A., Kong, M.: Computing the geodetic number of bipartite graphs. Congressus Numerantium pp. 113–120 (1995)
12. Gledel, V., Irsic, V., Klavzar, S.: Strong geodetic cores and cartesian product graphs. Appl. Math. Comput. **363** (2019)
13. Harary, F., Loukakis, E., Tsouros, C.: The geodetic number of a graph. Mathematical and Computer Modelling **17**(11), 89–95 (1993)
14. Iršič, V.: Strong geodetic number of complete bipartite graphs and of graphs with specified diameter. Graphs Comb. **34**(3), 443–456 (2018)
15. Iršič, V., Konvalinka, M.: Strong geodetic problem on complete multipartite graphs. Ars Math. Contemp. **17**(2), 481–491 (2019)
16. Klavzar, S., Manuel, P.D.: Strong geodetic problem in grid-like architectures. Bulletin of the Malaysian Mathematical Sciences Society **41**(3), 1671–1680 (2018)
17. Manuel, P., Klavžar, S., Xavier, A., Arokiaraj, A., Thomas, E.: Strong edge geodetic problem in networks. Open Mathematics **15**(1), 1225–1235 (2017)
18. Mezzini, M.: An o (mn2) algorithm for computing the strong geodetic number in outerplanar graphs. Discuss. Math. Graph Theory (2020)
19. Papadimitriou, C.H., Yannakakis, M.: Optimization, approximation, and complexity classes **43**(3), 425–440 (1991)
20. Rehmani, S., Sunitha, M.: Edge geodesic number of a fuzzy graph. Journal of Intelligent & Fuzzy Systems **37**(3), 4273–4286 (2019)
21. Santhakumaran, A., Ullas Chandran, S.: The edge geodetic number and cartesian product of graphs. Discussiones Mathematicae Graph Theory **30**(1), 55–73 (2010)
22. Santhakumaran, A., John, J.: Edge geodetic number of a graph. Journal of Discrete Mathematical Sciences and Cryptography **10**(3), 415–432 (2007)
23. Santhakumaran, A., Ullas Chandran, S.: Comment on" edge geodetic covers in graphs. Proyecciones (Antofagasta) **34**(4), 343–350 (2015)
24. Weller, M., Chateau, A., Giroudeau, R., König, J., Pollet, V.: On residual approximation in solution extension problems. J. Comb. Optim. **36**(4), 1195–1220 (2018)
25. Zmazek, E.: Strong edge geodetic problem on grids. Bulletin of the Malaysian Mathematical Sciences Society pp. 1–20 (2021)

Algorithm 2: Strong Geodetic Number Algorithm For Complete Multipartite Graphs

Data: L the list of the size of the parts of the complete multipartite graph G

Result: The strong geodetic number of G

```
n ← sum(L)
k ← len(L)
nc ← 0
for j in (0..L[0]) do
    d[0][j][L[0] − j] ← j ∗ (j − 1)/2
end
for i in (0..k − 2) do
    nc ← nc + L[i]
    for j in (0..nc) do
        for r in (0..nc) do
            if d[i][j][r] == −1 then
                continue
            end
            dc ← d[i][j][r]
            for s in (0..L[i + 1]) do
                for sp in (0..s ∗ (s − 1)/2) do
                    for dp in (0..dc) do
                        rp ← max(0, r − sp) + max(0, L[i + 1] − s − dp)
                        dd ← s ∗ (s − 1)/2 − sp + max(0, −(r − sp)) +
                          max(0, −(L[i + 1] − s − dp)) + dc − dp
                        if dd > d[i + 1][j + s][rp] then
                            d[i + 1][j + s][rp] = dd
                        end
                    end
                end
            end
        end
    end
end
j ← 0 ;
while d[k − 1][j][0] < 0 do
    j ← j + 1
end
return j;
```