Algorithms and Complexity for Geodetic Sets on Planar and Chordal Graphs

Dibyayan Chakraborty
Indraprastha Institute of Information Technology, Delhi, India

Sandip Das
Indian Statistical Institute, Kolkata, India

Florent Foucaud
Univ. Bordeaux, Bordeaux INP, CNRS, LaBRI, UMR5800, F-33400 Talence, France

Harmender Gahlawat
Indian Statistical Institute, Kolkata, India

Dimitri Lajou
Univ. Bordeaux, Bordeaux INP, CNRS, LaBRI, UMR5800, F-33400 Talence, France

Bodhayan Roy
Indian Institute of Technology, Kharagpur, India

Abstract

We study the complexity of finding the *geodetic number* on subclasses of planar graphs and chordal graphs. A set S of vertices of a graph G is a *geodetic set* if every vertex of G lies in a shortest path between some pair of vertices of S. The Minimum Geodetic Set (MGS) problem is to find a geodetic set with minimum cardinality of a given graph. The problem is known to remain NP-hard on bipartite graphs, chordal graphs, planar graphs and subcubic graphs. We first study MGS on restricted classes of planar graphs: we design a linear-time algorithm for MGS on solid grids, improving on a 3-approximation algorithm by Chakraborty et al. (CALDAM, 2020) and show that MGS remains NP-hard even for subcubic partial grids of arbitrary girth. This unifies some results in the literature. We then turn our attention to chordal graphs, showing that MGS is fixed parameter tractable for inputs of this class when parameterized by their *treewidth* (which equals the clique number minus one). This implies a linear-time algorithm for k-trees, for fixed k. Then, we show that MGS is NP-hard on interval graphs, thereby answering a question of Ekim et al. (LATIN, 2012). As interval graphs are very constrained, to prove the latter result we design a rather sophisticated reduction technique to work around their inherent linear structure.

2012 ACM Subject Classification Theory of computation → Design and analysis of algorithms

Keywords and phrases Geodetic set, Planar graph, Chordal graph, Interval graph, FPT algorithm

Digital Object Identifier 10.4230/LIPIcs.ISAAC.2020.7

Related Version A full version of the paper is available at `https://arxiv.org/abs/2006.16511`.

Funding This research was financed by the IFCAM project "Applications of graph homomorphisms" (MA/IFCAM/18/39) and the ANR project HOSIGRA (ANR-17-CE40-0022).
Bodhayan Roy: B. Roy is supported by an ISIRD Grant from Sponsored Research and Industrial Consultancy, IIT Kharagpur.

1 Introduction

A simple undirected graph G has vertex set $V(G)$ and edge set $E(G)$. For two vertices $u, v \in V(G)$, let $I(u, v)$ denote the set of all vertices in G that lie in some shortest path between u and v. For a subset S of vertices of a graph G, let $I(S) = \bigcup_{u,v \in S} I(u, v)$. We say that $T \subseteq V(G)$ is *covered* by S if $T \subseteq I(S)$. A set of vertices S is a *geodetic set* if

$V(G)$ is covered by S. The *geodetic number*, denoted $gn(G)$, is the minimum integer k such that G has a geodetic set of cardinality k. Given a graph G, the Minimum Geodetic Set (MGS) problem, introduced in [17], is to compute a geodetic set of G with minimum cardinality. In this paper, we study the computational complexity of MGS in subclasses of planar and chordal graphs. MGS is a natural graph covering problem that falls in the class of problems dealing with the important geometric notion of *convexity*: see [11, 22] for some general discussion of graph convexities. The setting of MGS is quite natural, and it can be applied to facility location problems such as the optimal determination of bus routes in a public transport network [6]. See also [10] for further applications.

The algorithmic complexity of MGS has been studied actively. In 1993, Harari, Loukakis and Tsouros, in [17], proved that MGS is NP-hard. Later, Dourado et al. [8, 9] strengthened the above result to *bipartite* graphs, *chordal* graphs (*i.e.* graphs with no induced cycle of order greater than 3) and *chordal bipartite* graphs (*i.e.* bipartite graphs with no induced cycle of order greater than 4). Recently, Bueno et al. [4] proved that MGS remains NP-hard for *subcubic* graphs, and Chakraborty et al. [6] proved that MGS is NP-hard for planar graphs. Kellerhals and Koana [19] studied the parameterized complexity of MGS, proving that it is unlikely to be FPT for the parameters solution size, feedback vertex set number and pathwidth, combined.

On the positive side, polynomial-time algorithms to solve MGS are known for *cographs* [8], *split* graphs [8], *ptolemaic* graphs [11], *block cactus* graphs [10], *outerplanar* graphs [21] and *proper interval* graphs [10], and the problem is FPT for parameters tree-depth and feedback edge set number [19].

A *grid embedding* of a graph is a set of points in two dimensions with integer coordinates such that each point in the set represents a vertex of the graph and, for each edge, the points corresponding to its endpoints are at Euclidean distance 1. A graph is a *partial grid* if it has a grid embedding. A graph is a *solid grid* if it has a grid embedding such that all interior faces have unit area. Chakraborty et al. [6] gave a 3-approximation algorithm for MGS on solid grids. We improve this as follows.

▶ **Theorem 1.** *There is a linear-time algorithm for MGS on solid grids.*

We note that researchers have designed polynomial-time algorithms for various problems on solid grids [12, 20, 24]. Our algorithm on solid grids does not require the grid embedding to be part of input. This is interesting since deciding whether an input graph is a solid grid is an NP-complete problem [16]. To complement Theorem 1, we prove the following.

▶ **Theorem 2.** *MGS is NP-hard for subcubic partial grids of girth at least g, for any fixed integer $g \geq 4$.*

We note that this result jointly strengthens three existing hardness results: for bipartite graphs [8], subcubic graphs [4] and planar graphs [6]. Moreover, partial grids are subclasses of many other important graph classes such as *disk* graphs, *rectangle intersection* graphs, etc [7, 23]. Hence, our result implies that MGS remains NP-hard on the aforementioned graph classes.

An *interval representation* of a graph G is a collection of intervals on the real line such that two intervals intersect if and only if the corresponding vertices are adjacent in G. A graph is an *interval graph* if it has an interval representation. Ekim et al. [10] asked if there is a polynomial-time algorithm for MGS on *interval graphs*. We give a negative answer to their question (note that proper interval graphs are those interval graphs with no induced $K_{1,3}$).

▶ **Theorem 3.** *MGS is NP-hard for interval graphs (even with no induced $K_{1,5}$).*

This result is somewhat surprising, as most covering problems can be solved in polynomial time on interval graphs (but other distance-based problems, like Metric Dimension, are NP-complete for interval graphs [13]). Our reduction (from 3-Sat) uses a quite involved novel technique, that we hope can be used to prove similar results for other distance-related problems on interval graphs. The main challenge here is to overcome the linear structure of the graph to transmit information across the graph. To this end, we use a sophisticated construction of many parallel *tracks*, *i.e.* shortest paths with intervals of (mostly) the same length spanning roughly the whole graph, and such that each track is shifted with respect to the previous one. Each track represents shortest paths that will be used by solution vertices from our variable and clause gadgets. In between the tracks, we are able to build our gadgets.

We remark that MGS admits a polynomial-time algorithm on proper interval graphs by a nontrivial dynamic programming scheme [10]. Problems known to be NP-complete on interval graphs but polynomial-time solvable on proper interval graphs are very rare; two examples known to us are Equitable Coloring [15] and Induced Subgraph Isomorphism [18].

To complement Theorem 3, we design an FPT algorithm for MGS on interval graphs when parameterized by its *treewidth* which equals its *clique number* ω minus one. Observe that interval graphs are also *chordal graphs*, *i.e.* graphs without induced cycles of order greater than 3. We use dynamic programming on tree decompositions to prove the following.

▶ **Theorem 4.** *MGS can be solved in time $2^{2^{O(\omega)}}n$ for chordal graphs and in time $2^{O(\omega)}n$ for interval graphs, where n is the order of the input graph.*

This result applies to the following setting. A *k-tree* is a graph formed by starting with a complete graph on $(k + 1)$ vertices and then repeatedly adding vertices by making each added vertex adjacent to exactly k neighbors forming a $(k + 1)$-clique. Allgeier [1] gave a polynomial-time algorithm to solve MGS on *maximal outerplanar* graphs, which is a subclass of 2-trees, and thus our algorithm generalizes this result (note that 2-trees are both chordal and planar). Since all k-trees are chordal graphs, MGS can be solved in time $2^{2^{O(k)}}n$ for k-trees of order n. Recall that this is unlikely to hold for *partial* k-trees (which are exactly the graphs of treewidth at most k) since MGS is $W[1]$-hard for parameter treewidth [19].

Structure of the paper. In Section 2, we describe the algorithm for solid grids. In Section 3, we present the algorithm for chordal graphs. In Section 4, we prove hardness for partial grids. In Section 5, we prove hardness for interval graphs. We conclude in Section 6. Due to space restrictions, some of the proofs are only sketched. The complete proof details can be found in the full version of this paper: see [5].

2 A linear-time algorithm for solid grids

We here give a linear-time algorithm for MGS on solid grids and prove Theorem 1. In the remainder of the section, G denotes a solid grid and $\mathcal{R}$ its grid embedding. A path P in G is a *corner path* if (i) no vertex of P is a cut-vertex, (ii) both end-vertices of P have degree 2 in G, and (iii) all other vertices of P have degree 3 in G. Chakraborty et al. [6] proved:

▶ **Lemma 5** ([6])**.** *Any geodetic set of G contains at least one vertex from each corner path.*

Any geodetic set of G contains all vertices of degree 0 or 1. We say that a vertex v of G is a *corner vertex* if v is an end-vertex of some corner path. All corner vertices can be found in linear time even if the grid embedding is not provided as an input [6].

▶ **Definition 6.** *We say that* $u_1, u_2, \ldots, u_k$ *forms a* corner sequence *if for each* $1 \leq i \leq k-1$*,*
1. *there is a corner path with* u_i *and* u_{i+1} *as endpoints, and*
2. *there is no corner vertex in the clockwise traversal of the boundary of* $\mathcal{R}$ *from* u_i *to* u_{i+1}*.*

A corner sequence is *maximal* if it is not a subsequence of any other corner sequence. For a corner sequence S, let $|S|$ denote the length of S.

▶ **Lemma 7.** *Let* $\mathcal{S}$ *be the set of all maximal corner sequences of* G*, and let* t *be the number of vertices of* G *with degree* 1*. Then,* $gn(G) \geq t + \sum_{S \in \mathcal{S}} \lfloor |S|/2 \rfloor$*.*

Proof. Any geodetic set of G contains all vertices of degree 1 and therefore $gn(G) \geq t$. Now, let X be any geodetic set of G and $S \in \mathcal{S}$ be an arbitrary maximal corner sequence. Assume that $u_1, u_2, \ldots, u_{|S|}$ forms the maximal subsequence S. Lemma 5 implies that for each $1 \leq j < |S|$, at least one vertex of the corner path between u_j and u_{j+1} must belong to X. Observe that two corner paths may have at most one corner vertex in common. Moreover, a corner vertex cannot be in three corner paths. Therefore, X must contain at least $\left\lfloor \frac{|S|}{2} \right\rfloor$ vertices. Now, let P be a corner path with endpoints a, b and P' be a corner path with endpoints a', b'. If a, b and a', b' are in two different maximal corner subsequences, then P and P' have no vertex in common. ◀

Due to space constraints, we only sketch the proof. Let $\mathcal{S}$ be the set of all maximal corner sequences of G. For a maximal corner sequence $S = u_1, u_2, \ldots, u_k$ let $f(S)$ denote the set $\{u_2, u_4 \ldots, u_{k-k'}\}$ where $k' = 0$ if k is even and $k' = 1$, otherwise. Observe that $|f(S)| = \left\lfloor \frac{k}{2} \right\rfloor$. Let V_1 be the set of all vertices of degree 1. Now consider the sets $V_2 = \cup_{S \in \mathcal{S}} f(S)$ and $D = V_1 \cup V_2$. Indeed, one can prove that D is a geodetic set of G and by Lemma 7, we are done.

3 An FPT algorithm for chordal graphs parameterized by clique number

We now sketch the main ideas for proving Theorem 4.

We give an FPT algorithm for chordal graphs parameterized by the clique number (which is also the treewidth plus 1). We explain how to improve the complexity in the case of interval graphs after the proof of the chordal case. Our algorithm performs dynamic programming on a *nice tree decomposition* of the input chordal graph. The main idea behind the algorithm is that the internal bags of the tree-decomposition (*i.e.* those who disconnect the tree into non-empty graphs) induce clique cutsets (cliques whose removal disconnects the graph). Then, for two vertices u, v all whose shortest paths go through some clique cutset X, their behaviour (with respect to MGS and X) can be described in terms of X only. This key observation will be enough to design our algorithm.

Nice tree decompositions are a well-known tool for designing dynamic programming algorithms for graphs of bounded treewidth. In our notation, the set of vertices of the graph associated to a node v of the tree, its bag, is denoted X_v. A nice tree decomposition of a chordal graph (see [2]) is a rooted tree decomposition where the bag of every node induces a clique. Each node belong to one of the following types. A node is a *join node* if it has exactly two children, on the same bags as the node. An *introduce node* has a unique child whose bag contains exactly one vertex less. A *forget node* has a unique child whose bag contains exactly one more vertex. A *leaf node* is a leaf of the tree whose bag is empty. The root node is a leaf node.

For a nice tree decomposition and a node v, we define $G_{\leq v}$ as the subgraph of G induced by the vertices of the subtree of the decomposition rooted at v. If $u \in V(G)$ and X is a clique, we say that u is *close* to a nonempty set of vertices $A \subseteq X$ with respect to X, if $d(u,x) = d_u$ when $x \in A$ and $d(u,x) = d_u + 1$ when $x \in X \setminus A$ (for some integer d_u). Intuitively, if X is clique cutset which creates two connected components G_1 and G_2 when removed and if $u \in V(G_1)$ then the shortest paths from u to some vertices of the other components "tends to" go through vertices of A.

Each maximal clique of a chordal graph will be associated to some node of the tree decomposition. Intuitively speaking, as in most tree-width based dynamic programming schemes, we need to show how the number of local solutions (*i.e.* in a bag of the tree) is bounded by a function of the maximal size of a bag, and how it can be computed from the information already computed for the node's children. To do this, we will need the following lemma, which deals with how shortest paths interact with clique cutsets.

▶ **Lemma 8.** *Let X be a clique cutset of a chordal graph G. Let u, v be two vertices of G such that all paths from u to v intersect X. Let A (resp. B) be a nonempty set of vertices of X such that u (resp. v) is close to A (resp. B) with respect to X. Then, a vertex x belongs to $I(u,v) \cap X$ (where $I(u,v)$ is the set of vertices of G covered by a shortest path from u to v) if and only if $x \in A \cap B$ or, $A \cap B = \varnothing$ and $x \in A \cup B$.*

Lemma 8 implies that to compute an optimal partial solution (*i.e.* a subset of vertices of $G_{\leq v}$) for a given bag X_v, it is sufficient to "guess" for which subsets A of X_v, there will exist (in the future solution that will be computed for ancestors of v) a vertex y which is close to A with respect to X_v. Thus, roughly speaking, it will be sufficient to index our solutions by types depending on what subsets A of X_v are required to satisfy this property.

More precisely, for each node v of our tree decomposition, we compute a table of partial solutions, indexed by *types*. For a node v, a type $\tau = (\tau^{int}, \tau^{ext}, \tau^{bag}, \tau^{cov})$ is an element of $\{0,1\}^{2^{X_v}} \times \{0,1\}^{2^{X_v}} \times 2^{X_v} \times 2^{X_v}$ where 2^{X_v} is the power set of X_v. We see τ^{int} and τ^{ext} as Boolean vectors indexed by the elements of 2^{X_v}, and τ^{bag} and τ^{bag}, as subsets of X_v. The table of partial solutions of the node v, denoted as $sol\,[v,\tau]$, will contain an optimal partial solution for X_v of the given type τ, computed using the partial solutions of the children of v. Our goal is to compute all such partial solutions $sol\,[v,\tau]$. The partial solution $sol\,[v,\tau]$, must follow some additional constraints that we detail below. Note that it is possible that no partial solution verify those constraints. In this case, $sol\,[v,\tau]$ is left empty.

For a node v of type τ and for a set $A \subseteq X_v$ of vertices, the Boolean $\tau^{int}[A]$ represents whether there is some $y \in sol\,[v,\tau]$ such that y is close to A with respect to X_v ("int" stands for "interior"). For $A \subseteq X_v$, the Boolean $\tau^{ext}[A]$ represents whether we need to add, at a later step of our algorithm, some vertex y such that y is close to A with respect to X_v. Here, y is a vertex that needs to be added later to the solution, in the upper part of the tree ("ext" stands for "exterior"). By Lemma 8, it is not necessary to keep track of all such vertices, as it is sufficient to record which subsets of vertices of X_v they are close to. This is a crucial property used to construct our solution: if there exists such a y, then we can cover some vertices of the subtree using this y, and Lemma 8 tells us exactly how.

The set τ^{cov} represents the vertices of X_v that we require to cover with $sol\,[v,\tau]$. Due to the existence of join nodes, we might want to cover the other vertices of X_v at a later step of the algorithm. The set τ^{bag} represents the vertices of X_v in $sol\,[v,\tau]$, and is essentially used to know which solutions of the children nodes of a join node can be merged.

To formalise the notion of types associated with a node we introduce the following definition which essentially asserts what our solution table for a given node must satisfy to be correct. To this end, we define a helper graph H_v^τ that simulates the vertices (using simplicial vertices) whose types are required to belong to the (future) solution.

For a node v, let $\mathcal{T}_v$ denote the set of all types of v. For a fixed τ let H_v^τ denote the graph obtained by adding a vertex S to $G_{\le v}$ whenever there is a set $S \subseteq X_v$ with $\tau^{ext}[S] = 1$, and making S adjacent to each $x \in S$. Let $S_v^\tau = \{S : S \subseteq X_v, \tau^{ext}[S] = 1\}$ denote the newly added vertices.

▶ **Definition 9.** *Let v be a node of T. A 4-tuple $\tau = (\tau^{int}, \tau^{ext}, \tau^{cov}, \tau^{bag})$ of $\mathcal{T}_v$ is a "type associated with v" if there exists a set $D \subseteq V(H_v^\tau)$ such that the following hold.*

(i) *$S_v^\tau \subseteq D$ and $\tau^{bag} = D \cap X_v$.*

(ii) *For a vertex $w \in (V(G_{\le v}) \setminus X_v) \cup \tau^{cov}$ there exists a pair $w_1, w_2 \in D$ such that $w \in I(w_1, w_2)$ and $w_1 \in D \setminus S_v^\tau$.*

(iii) *For a subset $A \subseteq X_v$, we have $\tau^{int}[A] = 1$ if and only if $D \cap V(G_{\le v})$ contains a vertex which is close to A with respect to X_v.*

Moreover, we shall say that the set $D \setminus S_v^\tau$ is a "certificate" for (v, τ).

The proof of Theorem 4 boils down to showing, by induction, that it is possible to construct certificates of minimum cardinality for each (valid) pair (v, τ) in a total of $2^{2^{O(\omega(G))}} n$ time. This is possible as the tree decomposition contains a $O(n)$ nodes and for each of them, computing our table for one particular node can be achieved in $2^{2^{O(\omega(G))}}$ time. For the root r, there is only one type τ_0 and therefore the minimum cardinality certificate for (r, τ_0) is an optimal geodetic set of G.

For interval graphs, the tree decomposition is a path decomposition. τ^{int} and τ^{ext} can be chosen in $\{0,1\}^{\mathcal{A}}$, where $\mathcal{A}$ is a set of size $O(\omega)$, reducing the time complexity to $2^{O(\omega)} n$.

4 Hardness for partial grids

We now prove Theorem 2. Let $\mathcal{PG}(3, g)$ denote the class of subcubic partial grids of girth at least g. Given a graph G, a subset $S \subseteq V(G)$ is a *vertex cover* of G if every edge in $E(G)$ has at least one end-vertex in S. The problem Minimum Vertex Cover is to compute a vertex cover of an input graph G with minimum cardinality. To prove Theorem 2, we reduce the NP-hard Minimum Vertex Cover problem on cubic planar graphs [14] to MGS on graphs in $\mathcal{PG}(3, g)$. We use a result of Valiant [25] which says that any planar graph G with maximum degree at most 4 has a drawing using $O(|V(G)|)$ area where vertices are represented as points on the integer grid, and edges are drawn as rectilinear curves on the integer grid.

Let $\mathcal{R}$ be an embedding of a cubic planar graph G as described above. One can ensure that the distance between two vertices is at least 100, and two parallel lines are at distance at least 100. (Any large constant would be sufficient). We call such an embedding a *good embedding* of G. A set S of vertices of a graph is an *edge geodetic set* if every edge lies in some shortest path between a pair of vertices in S. Note that an edge geodetic set is also a geodetic set (if there are no isolated vertices). We need the following lemma.

▶ **Lemma 10.** *Let H be a graph having a geodetic set S which is also an edge geodetic set. If H' denotes the graph obtained by replacing each edge of H with a path having $k \ge 0$ edges, then S is a geodetic set of H'.*

Proof. Let $w \in V(H')$ be a new vertex that was introduced while replacing an edge e of H with a path. Let $u_e, v_e \in S$ such that e belongs to a shortest path P between u_e and v_e. Let P' be the path obtained by replacing each edge of P by a path having k edges. Observe that P' is a shortest path between u_e and v_e in H'. Hence w belongs to a shortest path between u_e and v_e in H'. Thus S is a geodetic set of H'. ◀

Overview of the reduction. From a cubic planar graph G with a given good embedding, first we construct a planar graph $f_1(G)$ having maximum degree at most 6 and girth 4. We show that G has a vertex cover of size k if and only if $f_1(G)$ has a geodetic set of size $3|V(G)|+k$. Then, we construct a graph $f_2(G) \in \mathcal{PG}(3,42)$ such that the geodetic numbers of $f_2(G)$ and $f_1(G)$ are the same. When $g > 42$, we construct a graph $f_3(G) \in \mathcal{PG}(3,g)$ such that the geodetic numbers of $f_3(G)$ and $f_2(G)$ are the same.

Construction of $f_1(G)$. From a cubic planar graph G with a given good embedding $\mathcal{R}$, we construct a graph $f_1(G)$ as follows. Each vertex v of G will be replaced by a *vertex-gadget* G_v which is shown in Figure 1. The edges outside of the vertex-gadgets will depend on $\mathcal{R}$. We assume that the edges incident with any vertex v are labeled e_i^v with $0 \le i < 3$, in such a way that the numbering increases counterclockwise around v with respect to the embedding (thus the edge vw will have two labels: e_i^v and e_j^w, for some $i,j \in \{1,2,3\}$). Consider two vertices v and w that are adjacent in G, and let e_i^v and e_j^w be the two labels of edge vw in G. Add the edges (t_i^v, t_j^w), $(y_{i,i+1}^v, y_{j-1,j}^w)$ and $(y_{i-1,i}^v, y_{j+1,j}^w)$ (See Figure 1). All indices are taken modulo 3. There are no other edges in $f_1(G)$. Observe that $f_1(G)$ is planar, and has maximum degree at most 6 and girth 4. We have the following lemma whose proof is omitted due to space restrictions.

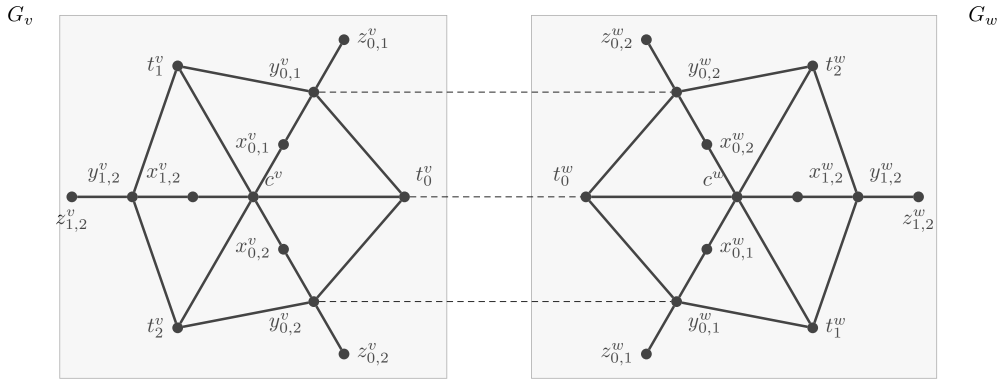

Figure 1 Illustration of vertex-gadgets in the reduction in the proof of Theorem 2. For an edge vw, we have the vertex gadgets G_v and G_w and the dashed lines indicate edges between two vertex-gadgets.

▶ **Lemma 11.** *The graph G has a vertex cover C of size k if and only if $f_1(G)$ has a geodetic set of size $3|V(G)|+k$.*

Proof. We construct a geodetic set S of $f_1(G)$ of size $3|V(G)|+k$ as follows. For each vertex v in G, we add the three vertices $z_{i,j}^v$ $(0 \le i < j \le 2)$ of G_v to S. If v is in C, we also add vertex c^v to S.

Let us show that S is indeed a geodetic set. First, we observe that in any vertex gadget G_v that is part of $f_1(G)$, the unique shortest path between two distinct vertices $z_{i,j}^v$, $z_{i',j'}^v$ has length 4 and goes through vertices $y_{i,j}^v$, t_k^v and $y_{i',j'}^v$ (where $\{k\} = \{i,j\} \cap \{i',j'\}$). Thus, it only remains to show that the vertices $\{c^v, x_{i,j}^v\}$ $(0 \le i < j \le 2)$ belong to some shortest path of vertices of S. Assume that v is a vertex of G in C. The shortest paths between c^v and $z_{i,j}^v$ have length 3 and one of them goes through vertex $x_{i,j}^v$. Thus, all vertices of G_v belong to some shortest path between vertices of S. Now, consider a vertex $w \notin C$ of G. Since G is a cubic planar graph, all three neighbours of w, say, w_1, w_2, w_3 must lie in C. Let $A = \{c^{w_1}, c^{w_2}, c^{w_3}\}$ and $Z = \{z_{0,1}^w, z_{1,2}^w, z_{0,2}^w\}$. Observe that all vertices of G_w lie in the set $I(A \cup Z)$. Therefore, S is a geodetic set.

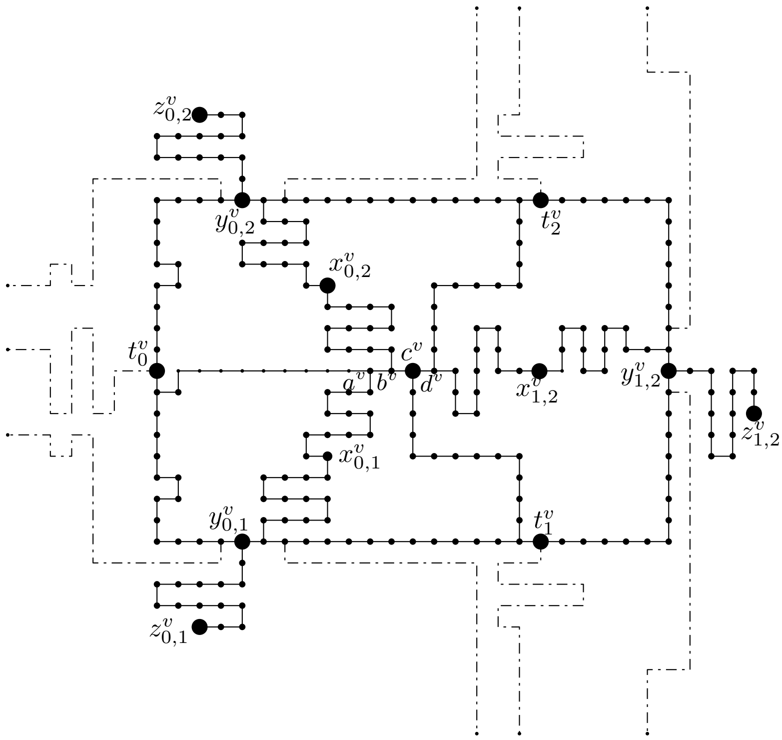

Figure 2 Construction of $f_2(G)$.

For the converse, assume we have a geodetic set S' of $f_1(G)$ of size $3|V(G)| + k$. We will show that G has a vertex cover of size k. First of all, observe that all the $3|V(G)|$ vertices of type $z^v_{i,j}$ are necessarily in S', since they have degree 1. As observed earlier, the shortest paths between those vertices already go through all vertices of type t^v_i and $y^v_{i,j}$. However, no other vertex lies on a shortest path between two such vertices: these shortest paths always go through the boundary 6-cycle of the vertex-gadgets. Let S'_0 be the set of the remaining k vertices of S'. These vertices are there to cover the vertices of type c^v and $x^v_{i,j}$. We construct a subset C' of $V(G)$ as follows: C' contains those vertices v of G whose vertex-gadget G_v contains a vertex of S'_0. We claim that C' is a vertex cover of G. Suppose by contradiction that there is an edge vw of G such that neither G_v nor G_w contains any vertex of S'_0. Without loss of generality assume that e^v_0 and e^w_0 are the two labels of the edge vw. Here also, the shortest paths between vertices of S always go through the boundary 6-cycle of G_v and thus, they never include vertex $x^v_{1,2}$. Let a and b be the neighbours of v different from w. Observe that no shortest path between a vertex of G_a and a vertex of G_b contains the vertex $x^v_{1,2}$, a contradiction. Thus S' is a vertex cover of G. ◀

Construction of $f_2(G)$. An edge uv of $f_1(G)$ is an *internal* edge if both u and v belong to G_w for some $w \in V(G)$. The other edges of $f_1(G)$ are *external* edges. We construct $f_2(G)$ in three steps described below.

1. Replace each vertex of type t_i^w (for $w \in V(G)$) with a new edge $T_i^w = (t_i^w t_i'^w)$. Replace each vertex of type $y_{i,j}^w$ with a path $Y_{i,j}^w = a_{i,j}^w\ y_{i,j}^w\ b_{i,j}^w\ d_{i,j}^w$. Replace each vertex of type c^w with a path $C^w = a^w\ b^w\ c^w\ d^w$. (See Figure 2).
2. Replace each internal edge between vertices having labels p, q with a new path such that the shortest path in the new graph between the vertices with label p, q has length 14 (this constant is required to get a valid good embedding).
3. For an edge $uv \in E(G)$, let E_{uv} denote the set of three external edges between G_u and G_v in $f_1(G)$. Recall that $\mathcal{R}$ is a good embedding of G. Let l_{uv} denote the length of the edge uv in $\mathcal{R}$. Replace all three external edges $p_i q_i \in E_{uv}$ $(1 \leq i \leq 3)$ with three new paths P_i $(1 \leq i \leq 3)$ such that lengths of all three paths are equal and in $O(l_{uv})$.

Clearly, $f_2(G_v)$ is a partial grid for each $v \in V(G)$ (Figure 2). It is not difficult to verify that $f_2(G)$ has maximum degree 3 and girth at least 42. Let C be a vertex cover of G with cardinality k. We construct a geodetic set S of $f_2(G)$ of cardinality $3|V(G)| + k$ as follows. For each vertex v in G, we add the three vertices with labels $z_{i,j}^v$ $(0 \leq i < j \leq 2)$ to S. If v is in C, we also add vertex c^v to S. From the construction of $f_2(G)$ and using arguments similar to that of Lemma 11, G has a vertex cover of size k if and only if $f_2(G)$ has a geodetic set of size $3|V(G)| + k$. Moreover, we can prove the following.

▶ **Lemma 12.** *The set S is both a geodetic set of minimum cardinality and an edge geodetic set of minimum cardinality of $f_2(G)$.*

Completion of the proof of Theorem 2. If $g \leq 42$, then observe that $f_2(G) \in \mathcal{PG}(3, g)$ and from the previous discussions, we have that MGS is NP-hard for graphs in $\mathcal{PG}(3, g)$. Otherwise, we replace every edge of $f_2(G)$ with a path of length g. Call this modified graph $f_3(G)$, and observe that $f_3(G) \in \mathcal{PG}(3, g)$. By Lemma 12, S is both a geodetic set of minimum cardinality and an edge geodetic set of minimum cardinality in $f_2(G)$. Now, due to Lemma 10, S is a geodetic set of $f_3(G)$ of cardinality $3|V(G)| + k$.

5 Hardness for interval graphs

We now give a sketch of the proof of Theorem 3. Let F be an instance of 3-Sat with variables $x_1, \dots, x_n$ and clauses $C_1, \dots, C_m$. We construct a set D of intervals in polynomial time such that the geodetic number of the intersection graph of D (denoted as $\mathcal{I}(D)$) is at most $4 + 7n + 58m$ if and only if F is a positive instance of 3-Sat.

The key intuition that explains why the problem is hard on interval graphs, is that considering two solution vertices x, y, the structure of the covered set $I(x, y)$ can be very complicated. Indeed, it can be that many vertices lying "in between" x and y in the interval representation, are not covered. This allows us to construct gadgets, by controlling which such vertices get covered, and which do not. Moreover, we can easily force some vertices to be part of the solution by representing them by an interval of length 0 (then, they are simplicial vertices), which is very useful to design our reduction. Nevertheless, implementing this idea turns out to be far from trivial, and to this end we need the crucial idea of *tracks*, which are shortest paths spanning a large part of the construction. Each track starts at a key interval called its *root* (representing a literal, for example) and serves as a shortest path from the root to the rightmost end of the construction. In a way, each track "carries the effect of the root" being chosen in a solution to the rest of the graph. The tracks are shifted in a way that no shortcut can be used going from one track to another. We are then able to locally modify the tracks and place our gadgets so that the track of, say, a literal, enables the interval of that literal to cover an interval of a specific clause gadget (while the other tracks are of no use for this purpose).

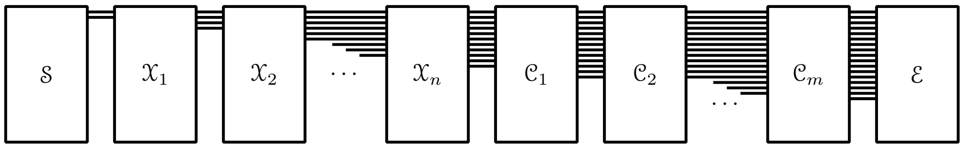

Figure 3 Overview of the construction. Boxes represent gadgets, and lines represent tracks.

Overview of the reduction. There are four stages of our reduction. We initialise it by constructing a set of intervals which we call the *start gadget* (denoted as $\mathcal{S}$). After creating the start gadget, we create the variable gadgets, which are placed consecutively, after the start gadget. For each variable x_i with $1 \leq i \leq n$, we create the variable gadget $\mathcal{X}_i$. Each variable gadget is composed of several *implication gadgets*. An implication gadget $IMP\left[\neg \boldsymbol{p} \rightarrow \boldsymbol{q}\right]$ ensures (under some extra hypotheses) that if $\boldsymbol{p}$ is not chosen in a geodetic set of our constructed intervals, then $\boldsymbol{q}$ must be chosen. These are used to encode the behaviour of the variables of the 3-SAT instance: there will be two possible solutions, corresponding to both truth values of x_i. After creating all the variable gadgets, we create the clause gadgets, also placed consecutively, after the variable gadgets. For each clause C_j with $1 \leq j \leq m$, we construct the clause gadget $\mathcal{C}_j$. Each clause gadget is composed of a *covering gadget*, several implication gadgets and several *AND gadgets*. The covering gadget of a clause C_i is denoted by $\text{COV}[i]$. For two intervals $\boldsymbol{p}$ and $\boldsymbol{q}$, the corresponding AND gadget is denoted by $\text{AND}\left[\boldsymbol{p}, \boldsymbol{q}\right]$. Together, these gadgets will ensure that all intervals of the clause gadget corresponding to the clause C_i are covered by six intervals if and only if one of the intervals corresponding to the literals of C_i is chosen in a geodetic set. This encodes the behaviour of the clauses of the 3-SAT instance.

After creating all the clause gadgets, we conclude our construction by creating the *end gadget* $\mathcal{E}$, placed after all clause gadgets. See Figure 3 for a schematic diagram.

Notations. Let S be a set of intervals with no isolated interval. For a vertex $v \in V(\mathcal{I}(S))$, let $\boldsymbol{v} = [min(v), max(v)]$ denote the interval corresponding to v in S, where $min(\boldsymbol{v})$ and $max(\boldsymbol{v})$ refer the left boundary and right boundary of $\boldsymbol{v}$, respectively. From now on, we only work with intervals. The *rightmost neighbour* of $\boldsymbol{v}$ is the interval intersecting $\boldsymbol{v}$ that has the maximum right boundary. For a nonempty set S of intervals, let $min(S) = \min\{min(\boldsymbol{v}) \colon \boldsymbol{v} \in S\}$, $max(S) = \max\{max(\boldsymbol{v}) \colon \boldsymbol{v} \in S\}$.. For two intervals $\boldsymbol{u}, \boldsymbol{v}$ we have $\boldsymbol{u} < \boldsymbol{v}$ if $max(\boldsymbol{u}) < min(\boldsymbol{v})$. Let S be a set of intervals and $\boldsymbol{u}, \boldsymbol{v} \in S$. A shortest path between $\boldsymbol{u}, \boldsymbol{v}$ is a shortest path between $\boldsymbol{u}, \boldsymbol{v}$ in $\mathcal{I}(S)$. The set $I(\boldsymbol{u}, \boldsymbol{v})$ is the set of intervals that belongs to some shortest path between $\boldsymbol{u}, \boldsymbol{v}$. The geodetic set of S is analogously defined. For a subset S' of S the phrase "S is covered by S'" means that S' is a geodetic set of S. A *point interval* is an interval of the form $[a, a]$. A *unit interval* is an interval of the form $[a, a+1]$. A set of intervals is *proper* if no interval contains another. A set $T = \{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_t}\}$ of intervals is a *track* if $max(\boldsymbol{u_i}) = \min\left(\boldsymbol{u_{i+1}}\right)$ for all $1 \leq i < t$ and no $\boldsymbol{u_i}$ is a point interval for all $1 \leq i \leq t$. Observe that if T is a track, then $\mathcal{I}(T)$ is a path. In our construction, each track T will be associated with a number of intervals called its *roots*. The set of roots of T is denoted by $R(T)$. For an intuition of how the tracks and roots are used, the track for which $\boldsymbol{v}$ is the root will almost entirely be a shortest path from $\boldsymbol{v}$ to any interval $\boldsymbol{w}$ to the right of $\boldsymbol{v}$ (except for some local shortcuts in the gadgets involving $\boldsymbol{w}$, that can be controlled).

▶ **Definition 13.** *Let T and T' be two tracks such that $T \cup T'$ is a proper set of intervals. Then $T < T'$ if $max(T) < max(T')$.*

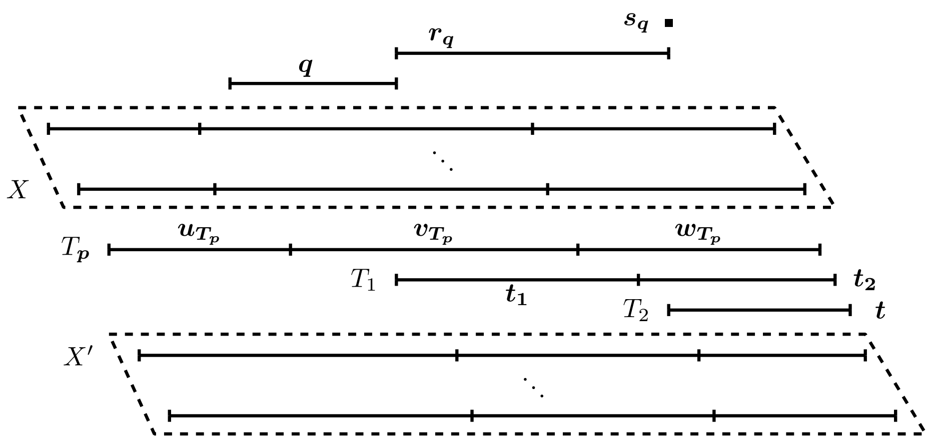

Figure 4 The implication gadget $IMP[\neg \boldsymbol{p} \to \boldsymbol{q}]$.

Let $\mathcal{T}$ be a set of tracks and $T \in \mathcal{T}$. By construction each $T \cup T'$ will be a proper set of intervals for $T, T' \in \mathcal{T}$. The phrase "*the track just preceding* T" refers to the track T' such that $T' < T$ and there is no T'' such that $T' < T'' < T$. The phrases "*the track just following* T", "*maximal track of* $\mathcal{T}$" and "*minimal track of* $\mathcal{T}$" are analogously defined.

Construction of $\mathcal{S}$. Let $\epsilon = \frac{1}{(n+m)^4}$ where n is the number of variables and m is the number of clauses. The start gadget $\mathcal{S}$ consists of four intervals which are defined as follows: the *start interval* $\boldsymbol{o} = [1,1]$, $\boldsymbol{u_o} = [1,2]$, the *true interval* $\top = [1+\epsilon, 1+\epsilon]$ and $\boldsymbol{u}_\top = [1+\epsilon, 2+\epsilon]$. Let $T_1 = \{\boldsymbol{u_o}\}$ and $T_2 = \{\boldsymbol{u}_\top\}$. Observe that T_1, T_2 are tracks and $T_1 < T_2$.

We initialize two more sets, the set $\mathcal{T} = \{T_1, T_2\}$ of all tracks, and the set $D = \mathcal{S}$ of all intervals. As we proceed with the construction, we shall insert more intervals in T_1, T_2 while maintaining that both of them are tracks. We shall also add more tracks in $\mathcal{T}$. Let $R(T_1) = \{\boldsymbol{o}\}$ and $R(T_2) = \{\top\}$.

Implication gadget of a root $\boldsymbol{p}$. To construct the variable gadgets and the clause gadgets, we need to define the implication gadget. On Figure 4, we present the implication gadget of a root $\boldsymbol{p}$ which is different from $\boldsymbol{o}$ of $\mathcal{S}$. The track $T_{\boldsymbol{p}} \in \mathcal{T}$ is the track such that $\boldsymbol{p} \in R(T_{\boldsymbol{p}})$. Since $\boldsymbol{p} \neq \boldsymbol{o}$, $T_{\boldsymbol{p}}$ is not the minimal element in $\mathcal{T}$. The interval $\boldsymbol{q}$ is a new interval constructed in the gadget for which we create a track T_1. The goal of this gadget is to ensure that $\boldsymbol{q}$ is part of our solution when $\boldsymbol{p}$ is not.

Construction of variable gadgets. We construct the variable gadgets sequentially and connect each of them to the previous one ($\mathcal{X}_1$ is connected to the start gadget $\mathcal{S}$). Assuming that we have placed $\mathcal{S}, \mathcal{X}_1, \ldots, \mathcal{X}_{i-1}$, we construct $\mathcal{X}_i$ as follows. For variable x_i, the gadget $\mathcal{X}_i$ consists of two implication gadgets. Let D and $\mathcal{T}$ be the set of intervals and tracks created so far. First, we construct $IMP[\neg\top \to \boldsymbol{x_i}]$. Observe that the sets D and $\mathcal{T}$ have been updated after the last operation. There is an interval $\boldsymbol{x_i}$ in D and there is a track $T \in \mathcal{T}$ whose root is $\boldsymbol{x_i}$. Now we construct $IMP[\neg\boldsymbol{x_i} \to \overline{\boldsymbol{x_i}}]$. Observe that after constructing all the variable gadgets, for each literal ℓ, there is an interval named $\boldsymbol{\ell}$ in D. Also note that nothing prevent us, at this point, from taking both $\boldsymbol{x_i}$ and $\overline{\boldsymbol{x_i}}$ in our solution. This property will follow from the cardinality constraints on the size of the solution.

A clause gadget consists of a covering gadget, several implication and AND gadgets.

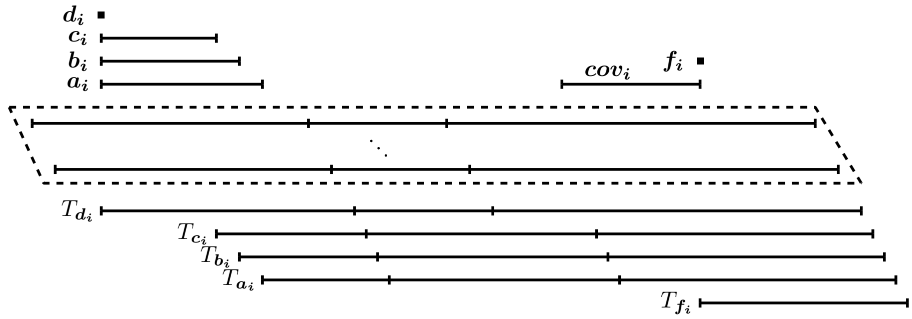

Figure 5 The covering gadget COV[i].

Construction of covering gadgets. The covering gadget COV$[i]$ of a clause C_i is presented on Figure 5. In particular, this gadget is used to say that, under some extra assumptions, one of $\boldsymbol{a_i}$, $\boldsymbol{b_i}$ and $\boldsymbol{c_i}$ is in our solution and covers the interval $\boldsymbol{cov_i}$ with a shortest path to $\boldsymbol{f_i}$.

Construction of the AND gadget. For two previously constructed interval $\boldsymbol{p}$ and $\boldsymbol{q}$, we can construct a gadget named AND $[\boldsymbol{p}, \boldsymbol{q}]$. We do not present the construction of this gadget as it is slightly more complicated than the previous ones. This gadget contains a number of new intervals. Among them is a particular interval denoted by $\gamma(p, q)$. The role of this gadget is as follows. If $\boldsymbol{p}$ and $\boldsymbol{q}$ are in our solution then every interval of the gadget is covered. Otherwise we need to add an interval of the gadget to cover it, namely the interval $\boldsymbol{\gamma\,(p, q)}$ is sufficient to cover every interval of AND $[\boldsymbol{p}, \boldsymbol{q}]$.

Construction of $\mathcal{C}_i$. We shall complete our construction of clause gadget $\mathcal{C}_i$ corresponding to the clause $C_i = (\ell_i^1, \ell_i^2, \ell_i^3)$. First, we create the covering gadget COV$[i]$ and update $D, \mathcal{T}$. Recall from the construction of COV$[i]$ that the intervals named $\boldsymbol{a_i}$, $\boldsymbol{b_i}$, $\boldsymbol{c_i}$ exist. Also recall from the construction of variable gadgets that the intervals $\boldsymbol{\ell_i^1}$, $\boldsymbol{\ell_i^2}$, $\boldsymbol{\ell_i^3}$ and $\overline{\boldsymbol{\ell_i^1}}$, $\overline{\boldsymbol{\ell_i^2}}$, $\overline{\boldsymbol{\ell_i^3}}$ exist. Now we create, in this order, $IMP\left[\neg \boldsymbol{a_i} \rightarrow \boldsymbol{a_i'}\right]$, AND $\left[\boldsymbol{a_i}, \boldsymbol{\ell_i^1}\right]$, AND $\left[\boldsymbol{a_i'}, \overline{\boldsymbol{\ell_i^1}}\right]$, $IMP\left[\neg \boldsymbol{b_i} \rightarrow \boldsymbol{b_i'}\right]$, AND $\left[\boldsymbol{b_i}, \boldsymbol{\ell_i^2}\right]$, AND $\left[\boldsymbol{b_i'}, \overline{\boldsymbol{\ell_i^2}}\right]$, $IMP\left[\neg \boldsymbol{c_i} \rightarrow \boldsymbol{c_i'}\right]$, AND $\left[\boldsymbol{c_i}, \boldsymbol{\ell_i^3}\right]$, AND $\left[\boldsymbol{c_i'}, \overline{\boldsymbol{\ell_i^3}}\right]$ where $\boldsymbol{a_i'}$, $\boldsymbol{b_i'}$ and $\boldsymbol{b_i'}$ are three new intervals constructed in the corresponding implication gadgets. The role of this last part is to ensure that $\boldsymbol{a_i}$ and $\boldsymbol{\ell_i^1}$ are either both in the solution or both not in the solution. The reality is slightly more complex but one should think of $\boldsymbol{a_i}$ as a copy of $\boldsymbol{\ell_i^1}$ and $\boldsymbol{a_i'}$ as a copy of $\overline{\boldsymbol{\ell_i^1}}$. The same holds for $\boldsymbol{b_i}$ and $\boldsymbol{c_i}$. This completes the construction of $\mathcal{C}_i$.

Construction of end gadget. For each $T \in \mathcal{T}$, we introduce two new intervals $\boldsymbol{u_T} = [max(T), max(T) + 1], \boldsymbol{e_T} = [max(\boldsymbol{u_T}), max(\boldsymbol{u_T})]$ and define $T = T \cup \{\boldsymbol{u_T}\}, D = D \cup \{\boldsymbol{u_T}, \boldsymbol{e_T}\}_{T \in \mathcal{T}}$. For each $T \in \mathcal{T}$, let $\boldsymbol{e_T}$ be the *tail* of T. The end gadget $\mathcal{E}$ consists of all the new intervals created above. The role of this gadget is to ensure that every interval belonging to a track is covered.

5.1 Analysis

First, note that there are $2 + 4n + 35m$ tracks in $\mathcal{T}$ and $n_{point} = 4 + 6n + 52m$ point intervals in D. The total number of intervals in D is $O((n + m)^2)$. Remark that the point intervals are exactly the simplicial vertices of D, hence they belong to every geodetic set of D. Let T be a track in $\mathcal{T}$ such that $T = \{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$ with $\max(\boldsymbol{u_i}) = \min(\boldsymbol{u_{i+1}})$ $(1 \leq i \leq k - 1)$. Observe that for each i with $1 \leq i \leq k - 1$, $\boldsymbol{u_{i+1}}$ is the rightmost neighbour of $\boldsymbol{u_i}$.

▶ **Proposition 14.** *Let* $\boldsymbol{u}$ *and* $\boldsymbol{v}$ *be two intervals of* D *such that* $\min(\boldsymbol{u}) < \min(\boldsymbol{v})$*. The path* $\boldsymbol{u_0}, \boldsymbol{u_1}, \ldots, \boldsymbol{u_k}, \boldsymbol{v}$ *is a shortest path from* $\boldsymbol{u}$ *to* $\boldsymbol{v}$ *(where* $\boldsymbol{u} = \boldsymbol{u_0}$*,* $\boldsymbol{u_{i+1}}$ *is the rightmost neighbour of* $\boldsymbol{u_i}$ *for* $i \in 1 \leq i \leq k-1$*, and* $\boldsymbol{u_{k-1}} \notin N(\boldsymbol{v})$*, while* $\boldsymbol{u_k} \in N(\boldsymbol{v})$*). We say that such a path is a* good shortest path.

An interval $\boldsymbol{u}$ is a *track interval* if $\boldsymbol{u} \in T$ for some $T \in \mathcal{T}$. From now on, U shall denote the set of track intervals. Let S_p be the set of all point intervals and recall that S_p is a subset of every geodetic set of D. For an interval $\boldsymbol{z}$, let $T(z)$ denote the track T such that $\boldsymbol{z} \in R(T)$. From our construction, one can observe that all track intervals are covered by pair of vertices in S_p.

▶ **Proposition 15.** *For an implication gadget* $IMP[\neg \boldsymbol{p} \to \boldsymbol{q}]$*, let* T *be the track with root* $\boldsymbol{q}$*. Then* $\boldsymbol{q} \in I(\boldsymbol{p}, \boldsymbol{s_q})$*,* $\boldsymbol{r_q} \in I(\boldsymbol{o}, \boldsymbol{s_q})$*.*

▶ **Proposition 16.** *Consider the cover gadget* $COV[i]$ *and let* $\boldsymbol{z} \in \{\boldsymbol{a_i}, \boldsymbol{d_i}, \boldsymbol{c_i}\}$*. Then* $\boldsymbol{z} \in I(\boldsymbol{d_i}, \boldsymbol{e_{T(z)}})$ *and* $\boldsymbol{cov_i} \in I(\boldsymbol{z}, \boldsymbol{f_i})$*.*

▶ **Proposition 17.** *Consider an AND gadget* $AND[\boldsymbol{p}, \boldsymbol{q}]$*. The set* $\{\boldsymbol{p}, \boldsymbol{q}\} \cup S_p$ *covers all vertices in* $AND[\boldsymbol{p}, \boldsymbol{q}]$*. The set* $\{\boldsymbol{\gamma}\left(\boldsymbol{p}, \boldsymbol{q}\right)\} \cup S_p$ *covers all vertices in* $AND[\boldsymbol{p}, \boldsymbol{q}]$ *where* $\boldsymbol{\gamma}\left(\boldsymbol{p}, \boldsymbol{q}\right)$ *is a vertex of* $AND[\boldsymbol{p}, \boldsymbol{q}]$*.*

We shall show that if F is satisfiable, then D has a geodetic set of cardinality $4+7n+58m = n_{point} + n + 6m$. Let $\phi\colon \{x_1, x_2, \ldots, x_n\} \to \{0, 1\}$ be a satisfying assignment of F (we also define $\phi(\overline{x_i}) = 1 - \phi(x_i)$). Now, define the following sets. Let $S_1 = \{\boldsymbol{x_i}\colon \phi(x_i) = 1\} \cup \{\overline{\boldsymbol{x_i}}\colon \phi(\overline{x_i}) = 1\}$. Let $S_2 = \emptyset$. Now, for each clause $C_i = (\ell_i^1, \ell_i^2, \ell_i^3)$, and for each $(\boldsymbol{v}, \boldsymbol{v'}, \boldsymbol{\ell}, \bar{\boldsymbol{\ell}}) \in \left\{(\boldsymbol{a_i}, \boldsymbol{a_i'}, \boldsymbol{\ell_i^1}, \overline{\boldsymbol{\ell_i^1}}), (\boldsymbol{b_i}, \boldsymbol{b_i'}, \boldsymbol{\ell_i^2}, \overline{\boldsymbol{\ell_i^2}}), (\boldsymbol{c_i}, \boldsymbol{c_i'}, \boldsymbol{\ell_i^3}, \overline{\boldsymbol{\ell_i^3}})\right\}$, if $\phi(\ell) = 1$, then put $S_2 = S_2 \cup \left\{\boldsymbol{v}, \boldsymbol{\gamma}\left(\boldsymbol{v'}, \bar{\boldsymbol{\ell}}\right)\right\}$ otherwise put $S_2 = S_2 \cup \{\boldsymbol{v'}, \boldsymbol{\gamma}\left(\boldsymbol{v}, \boldsymbol{\ell}\right)\}$. We have that $|S_1 \cup S_2 \cup S_p| = 4 + 7n + 58m$.

▶ **Lemma 18.** *The set* S *is a geodetic set of* D*.*

Proof. As $S_p \subseteq S$, we know that all track intervals of D are covered. Moreover, every interval of the form $\boldsymbol{r_q}$ is covered by Proposition 15. Consider any variable gadget $\mathcal{X}_i$ corresponding to the variable x_i. Recall from construction, that $\mathcal{X}_i = IMP[\neg\top \to \boldsymbol{x_i}] \cup IMP[\neg \boldsymbol{x_i} \to \overline{\boldsymbol{x_i}}]$. Due to Proposition 15, we have that $\boldsymbol{x_i} \in I(\top, \boldsymbol{e_{T(x_i)}})$. Hence, all intervals of $IMP[\neg\top \to \boldsymbol{x_i}]$ are covered by S. Due to Proposition 15, either $\overline{\boldsymbol{x_i}} \in I(\boldsymbol{x_i}, \boldsymbol{e_{T(\overline{x_i})}})$ when $\boldsymbol{x_i} \in S$ or $\overline{\boldsymbol{x_i}} \in S$ otherwise. The above arguments imply that all intervals in $\mathcal{X}_i$ are covered by S.

Now, consider any clause $C_i = (\ell_i^1, \ell_i^2, \ell_i^3)$ and recall the construction of $\mathcal{C}_i$. Observe that there exists at least one interval $\boldsymbol{z} \in \{\boldsymbol{a_i}, \boldsymbol{b_i}, \boldsymbol{c_i}\} \cap S$. Using Proposition 16, we can infer that all intervals in COV$[i]$ are covered by S. Now, consider the implication gadget $IMP\left[\neg \boldsymbol{a_i} \to \boldsymbol{a_i'}\right]$, note that $\left\{\boldsymbol{a_i}, \boldsymbol{a_i'}\right\} \cap S \neq \emptyset$ and therefore using Proposition 15, we can infer that all intervals in $IMP\left[\neg \boldsymbol{a_i} \to \boldsymbol{a_i'}\right]$ are covered by S. Repeating the above arguments for $IMP\left[\neg \boldsymbol{b_i} \to \boldsymbol{b_i'}\right]$ and $IMP\left[\neg \boldsymbol{c_i} \to \boldsymbol{c_i'}\right]$, we infer that all intervals in these implication gadgets are covered by S. Now, consider the AND gadget AND $\left[\boldsymbol{a_i}, \boldsymbol{\ell_i^1}\right]$. From our definition of S_2, it follows that either $\left\{\boldsymbol{a_i}, \boldsymbol{\ell_i^1}\right\} \subseteq S$ or $\boldsymbol{\gamma}\left(\boldsymbol{p}, \boldsymbol{q}\right) \in S$. In both cases, we can use Proposition 17 to show that all intervals in AND $\left[\boldsymbol{a_i}, \boldsymbol{\ell_i^1}\right]$ are covered by S. Repeating the above arguments for all the AND gadgets in $\mathcal{C}_i$, we can show that all intervals of $\mathcal{C}_i$ are covered by S. ◀

Now, we show that if the geodetic number of D is at most $4 + 7n + 58m$, then F is satisfiable. The next proposition is key in showing this (recall that U contains all track vertices). It is proven by considering an interval $\boldsymbol{v}$ with minimum $\min(\boldsymbol{v})$ contradicting the statement.

▶ **Proposition 19.** *There is a minimum-size geodetic set S^* of D such that $S^* \cap U = \emptyset$.*

A *good geodetic set* of D is a geodetic set of minimum cardinality which does not contain any interval belonging to a track (*i.e.* intervals of U).

▶ **Proposition 20.** *Let S^* be a good geodetic set of D and $IMP[\neg \boldsymbol{p} \to \boldsymbol{q}]$ be an implication gadget where $\boldsymbol{p}$ is the only root of $T(\boldsymbol{p})$. Then, either $\boldsymbol{p} \in S^*$ or $\boldsymbol{q} \in S^*$.*

▶ **Proposition 21.** *Let S^* be a good geodetic set of D and let $C_i = (\ell_i^1, \ell_i^2, \ell_i^3)$ be a clause. Then $|S^* \cap \mathcal{C}_i| \geq 6$. Moreover if none of $\boldsymbol{\ell_1^i}, \boldsymbol{\ell_2^i}, \boldsymbol{\ell_3^i}$ is in S^* then $|S^* \cap \mathcal{C}_i| \geq 7$.*

▶ **Lemma 22.** *If there is a good geodetic set of D with cardinality $4 + 7n + 58m$, then F is satisfiable.*

Proof. Let S^* be a good geodetic set of D with cardinality $4 + 7n + 58m$. Recall that a variable gadget $\mathcal{X}_i$ is $IMP[\neg\top \to \boldsymbol{x_i}] \cup IMP[\neg \boldsymbol{x_i} \to \overline{\boldsymbol{x_i}}]$. Due to Proposition 20, we know that at least one among $\{\boldsymbol{x_i}, \overline{\boldsymbol{x_i}}\}$ lies in S^*. Let $S_1 = (S^* \setminus S_p) \cap (\cup_{1 \leq i \leq n} \mathcal{X}_i)$, and $S_2 = (S^* \setminus S_p) \cap (\cup_{1 \leq i \leq m} \mathcal{C}_i)$. Note that $S_1 \cup S_2 \cup S_p \subseteq S^*$. We have $|S_1| \geq n$, $|S_3| \geq 6m$ by Proposition 21, and $|S_p| = 4 + 6n + 52m$. Therefore, $|S_1| = n$ as $|S^*| \leq 4 + 7n + 58m$. This means that for each $1 \leq i \leq n$, exactly one of $\boldsymbol{x_i}, \overline{\boldsymbol{x_i}}$ lies in S^*. Based on these, we define the following truth assignment ϕ: $\{x_1, \ldots, x_n\} \to \{1, 0\}$ of F. Define $\phi(x_i) = 1$ if $\boldsymbol{x_i} \in S^*$ and $\phi(x_i) = 0$, otherwise. Using Proposition 21 we can infer that for each $1 \leq i \leq m$, we have that $|S^* \cap \mathcal{C}_i| = 6$ and at least one of the intervals $\boldsymbol{\ell_1^i}, \boldsymbol{\ell_2^i}, \boldsymbol{\ell_3^i}$ lies in S^*. Thus, for at least one literal ℓ_i^j, we have that $\phi(\ell_i^j) = 1$, as needed. ◀

6 Conclusion

We gave a polynomial-time algorithm for MGS on solid grids and proved that MGS is NP-hard on partial grids and interval graphs. We proved that MGS is FPT on chordal graphs when parameterized by the clique number.

Are there FPT algorithms for MGS on interval graphs, chordal graphs, partial grids, planar graphs when parameterized by the geodetic number?

Assuming the *Exponential Time Hypothesis*, our reduction implies that there cannot be a $2^{o(\sqrt{n})}$ time algorithm for MGS on interval graphs of order n. Are there subexponential time algorithms for MGS on interval graphs or chordal graphs, matching this lower bound? (This is the case for many graph problems for geometric intersection graphs, see [3].)

We have seen that for every k, MGS is solvable in time $f(k)n$ for k-trees, but such a running time is unlikely to be possible for *partial* k-trees, since MGS is known to be W[1]-hard for parameter tree-width [19]. However, there could still exist an XP-time algorithm for MGS, running in time $n^{g(k)}$ on partial k-trees. In fact, it is already unknown whether MGS is solvable in polynomial time on partial 2-trees (also known as series-parallel graphs or K_4-minor-free graphs).

References

1 B. Allgeier. Structure and properties of maximal outerplanar graphs. *PhD Thesis, University of Louisville*, 2009.

2 R. Belmonte, P. A. Golovach, P. Heggernes, P. van't Hof, M. Kamiński, and D. Paulusma. Detecting fixed patterns in chordal graphs in polynomial time. *Algorithmica*, 69:501–521, 2014.

3 É. Bonnet and P. Rzążewski. Optimality program in segment and string graphs. *Algorithmica*, 81(7):3047–3073, 2019.

4 L. R. Bueno, L. D. Penso, F. Protti, V. R. Ramos, D. Rautenbach, and U. S. Souza. On the hardness of finding the geodetic number of a subcubic graph. *Information Processing Letters*, 135:22–27, 2018.

5 D. Chakraborty, S. Das, F. Foucaud, H. Gahlawat, D. Lajou, and B. Roy. Algorithms and complexity for geodetic sets on planar and chordal graphs, 2020. arXiv:2006.16511.

6 D. Chakraborty, F. Foucaud, H. Gahlawat, S. K. Ghosh, and B. Roy. Hardness and approximation for the geodetic set problem in some graph classes. In *CALDAM'20*, pages 102–115. Springer, 2020.

7 B. N. Clark, C. J. Colbourn, and D. S. Johnson. Unit disk graphs. *Discrete mathematics*, 86(1-3):165–177, 1990.

8 M. C. Dourado, F. Protti, D. Rautenbach, and J.L. Szwarcfiter. Some remarks on the geodetic number of a graph. *Discrete Mathematics*, 310(4):832–837, 2010.

9 M. C. Dourado, F. Protti, and J.L. Szwarcfiter. On the complexity of the geodetic and convexity numbers of a graph. In *ICDM*, volume 7, pages 101–108. Ramanujan Mathematical Society, 2008.

10 T. Ekim, A. Erey, P. Heggernes, P. van't Hof, and D. Meister. Computing minimum geodetic sets of proper interval graphs. In *LATIN'12*, pages 279–290. Springer, 2012.

11 M. Farber and R. E. Jamison. Convexity in graphs and hypergraphs. *SIAM Journal on Algebraic Discrete Methods*, 7(3):433–444, 1986.

12 A. E. Feldmann and P. Widmayer. An $o(n^4)$ time algorithm to compute the bisection width of solid grid graphs. In *European Symposium on Algorithms*, pages 143–154. Springer, 2011.

13 F. Foucaud, G. B. Mertzios, R. Naserasr, A. Parreau, and P. Valicov. Identification, location-domination and metric dimension on interval and permutation graphs. II. Algorithms and complexity. *Algorithmica*, 78:914–944, 2017.

14 M. R. Garey and D. S. Johnson. *Computers and intractability*, volume 29. W. H.Freeman New York, 2002.

15 G. C. M. Gomes, C. V. G. C. Lima, and V. F. dos Santos. Parameterized Complexity of Equitable Coloring. *Discrete Mathematics & Theoretical Computer Science*, vol. 21 no. 1, ICGT 2018, May 2019.

16 A. Gregori. Unit-length embedding of binary trees on a square grid. *Information Processing Letters*, 31(4):167–173, 1989.

17 F. Harary, E. Loukakis, and C. Tsouros. The geodetic number of a graph. *Mathematical and Computer Modelling*, 17(11):89–95, 1993.

18 P. Heggernes, P. van't Hof, D. Meister, and Y. Villanger. Induced subgraph isomorphism on proper interval and bipartite permutation graphs. *Theoretical Computer Science*, 562:252–269, 2015.

19 L. Kellerhals and T. Koana. Parameterized complexity of geodetic set. *arXiv*, 2020. arXiv:2001.03098.

20 F. Keshavarz-Kohjerdi, A. Bagheri, and A. Asgharian-Sardroud. A linear-time algorithm for the longest path problem in rectangular grid graphs. *Discrete Applied Mathematics*, 160(3):210–217, 2012.

21 M. Mezzini. Polynomial time algorithm for computing a minimum geodetic set in outerplanar graphs. *Theoretical Computer Science*, 745:63–74, 2018.

22 I. M. Pelayo. *Geodesic Convexity in Graphs.* Springer, 2013.

23 C. Thomassen. Interval representations of planar graphs. *Journal of Combinatorial Theory, Series B*, 40(1):9–20, 1986.

24 C. Umans and W. Lenhart. Hamiltonian cycles in solid grid graphs. In *Proceedings 38th Annual Symposium on Foundations of Computer Science*, pages 496–505. IEEE, 1997.

25 L. G. Valiant. Universality considerations in VLSI circuits. *IEEE Transactions on Computers*, 100(2):135–140, 1981.