Journal of Graph Algorithms and Applications
http://jgaa.info/ vol. 26, no. 4, pp. 401–419 (2022)
DOI: 10.7155/jgaa.00601

Parameterized Complexity of Geodetic Set

Leon Kellerhals[1] *Tomohiro Koana*[1]

[1]TU Berlin, Faculty IV, Algorithmics and Computational Complexity

Submitted: October 2021	Reviewed: April 2022	Revised: June 2022
Accepted: June 2022	Final: June 2022	Published: July 2022
Article type: Regular paper		Communicated by: Y. Okamoto

Abstract. A vertex set S of a graph G is *geodetic* if every vertex of G lies on a shortest path between two vertices in S. Given a graph G and $k \in \mathbb{N}$, the NP-hard Geodetic Set problem asks whether there is a geodetic set of size at most k. Complementing various works on Geodetic Set restricted to special graph classes, we initiate a parameterized complexity study of Geodetic Set and show, on the one side, that Geodetic Set is W[1]-hard when parameterized by feedback vertex number, path-width, and solution size, combined. On the other side, we develop fixed-parameter algorithms with respect to the feedback edge number, the tree-depth, and the modular-width of the input graph.

1 Introduction

Let G be an undirected, simple graph with vertex set $V(G)$ and edge set $E(G)$. The *interval* $I[u, v]$ of two vertices u and v of G is the set of vertices of G that are contained in any shortest path between u and v. In particular, $u, v \in I[u, v]$. For a set S of vertices, let $I[S]$ be the union of the intervals $I[u, v]$ over all pairs of vertices u and v in S. A set of vertices S is called *geodetic* if $I[S]$ contains all vertices of G. In this work we study the following problem (see an exemplary illustration in Figure 1):

Geodetic Set
Input: A graph G and an integer k.
Question: Does G have a geodetic set of cardinality at most k?

Atici [2] showed that Geodetic Set is NP-complete on general graphs, and it was shown that the hardness holds even if the graph is planar [8], subcubic [7], chordal, or bipartite chordal

TK was partially supported by the DFG projects FPTinP (NI 369/16) and MATE (NI 369/17). An extended abstract of this work appeared in the proceedings of the 15th International Symposium on Parameterized and Exact Computation (IPEC '20), held in Hong Kong, China (virtual conference), December 14–18, 2020.

E-mail addresses: leon.kellerhals@tu-berlin.de (Leon Kellerhals) tomohiro.koana@tu-berlin.de (Tomohiro Koana)

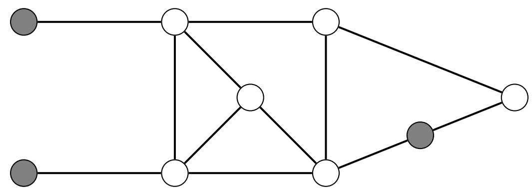

Figure 1: An exemplary graph. The gray vertices form a minimum geodetic set. The shortest paths between the top left and the bottom right gray vertex cover all vertices except for the bottom left vertex. Observe that every geodetic set contains all degree-one vertices.

[12]. Although not stated, W[2]-hardness for the solution size k directly follows from the reduction for the latter result of Dourado et al. [12]. On the positive side, the problem was shown to be polynomial-time solvable for cographs, split graphs and unit interval graphs [12]. Also, upper bounds on the geodetic set size in Cartesian product graphs were studied [6].

For a graph G and $k \in \mathbb{N}$, the closely related Geodetic Hull problem asks whether there is a vertex set $S \subseteq V(G)$ with $I^{|V(G)|}[S] = V(G)$ and $|S| \leq k$, where $I^0[S] = S$ and $I^j[S] = I[I^{j-1}[S]]$ for $j > 0$. Geodetic Hull is NP-hard on bipartite [1], chordal [4], and P_9-free graphs [13]. Recently, Kanté et al. [18] studied the *parameterized complexity* of Geodetic Hull: they proved that the problem is W[2]-hard when parameterized by k, and W[1]-hard but in XP when parameterized by tree-width.[1]

Our Contributions. Comparing the algorithmic complexity of Geodetic Hull and Geodetic Set, one can observe that both problems are trivial on trees (take all leaves into the solution). But while Geodetic Hull is polynomial-time solvable on graphs of constant tree-width, the complexity of Geodetic Set on graphs of tree-width two is unknown to the best of our knowledge. Motivated by this gap, we study the parameterized complexity of Geodetic Set for structural parameters such as tree-width that measure the tree-likeness of the input graph.

We start off by showing that Geodetic Set is W[1]-hard with respect to tree-width. More specifically, we show that Geodetic Set is W[1]-hard for feedback vertex number, path-width, and solution size, all three combined (Section 3), using a parameterized reduction from the W[1]-hard Grid Tiling problem [21]. Since this reduction implies NP-hardness, this complements previous results by providing a more fine-grained view on computational tractability in terms of parameterized complexity instead of studying special graph classes.

We complement the W[1]-hardness by presenting two fixed-parameter tractability results for Geodetic Set. First, we show that Geodetic Set is fixed-parameter tractable with respect to the feedback edge number (Section 4). It turns out to be quite effortful to obtain fixed-parameter tractability, requiring the design and analysis of polynomial-time data reduction rules and branching before employing Integer Linear Programming (ILP) with a bounded number of variables to determine the final positions of the solution vertices. To the best of our knowledge, this is the first usage of ILP when solving Geodetic Set.

Second, we show that Geodetic Set is fixed-parameter tractable with respect to clique-width combined with diameter (Section 5); note that Geodetic Set is NP-hard even on graphs with constant diameter [12], and W[1]-hard with respect to clique-width (this follows from our first result). Our result exploits the fact that we can express Geodetic Set in an MSO_1 logic formula, the

[1] Informally, this means it can be solved in polynomial time for graphs of constant tree-width.

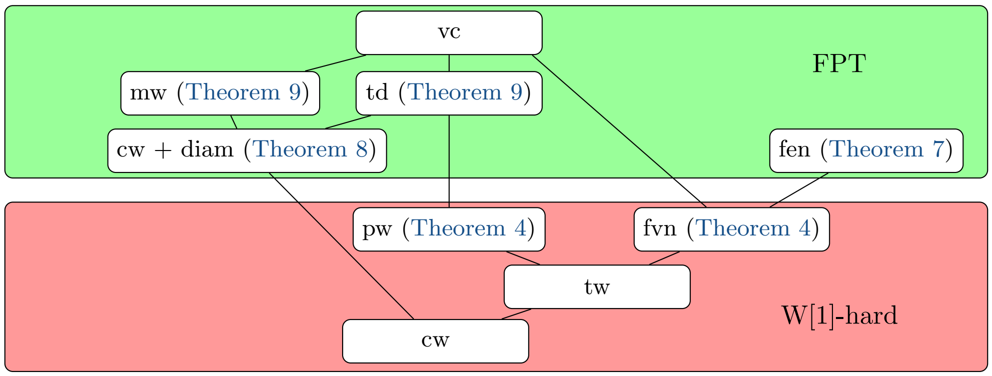

Figure 2: An overview of our results for Geodetic Set, containing the parameters vertex cover number (vc), modular-width (mw), tree-depth (td), clique-width (cw), diameter (diam), feedback edge number (fen), path-width (pw), feedback vertex number (fvn) and tree-width (tw). An edge between two parameters indicates that the one below is smaller than some function of the other.

length of which is upper-bounded in a function of the diameter of the graph. A direct consequence of this result is that Geodetic Set is fixed-parameter tractable with respect to tree-depth and with respect to modular-width.

Figure 2 gives an overview of the parameters for which we obtain positive and negative results, and presents their interdependence.

2 Preliminaries

For $n \in \mathbb{N}$ let $[n] = \{1, 2, \ldots, n\}$. The *distance* $d_G(u, v)$ between two vertices u and v in G is the length of a shortest path between u and v (also called *shortest* u–v*-path*). We drop the subscript $\cdot_G$ if G is clear from context. Note that w belongs to $I[u, v]$ if and only if $d_G(u, v) = d_G(u, w) + d_G(w, v)$. The *diameter* $\mathrm{diam}(G)$ of G is the maximum distance between any two vertices of G. A *multigraph* G consists of a vertex set and an edge multiset. Note that in a multigraph, we count self-loops twice for the vertex degree.

A set $F \subseteq E(G)$ is a *feedback edge set* if $G \setminus F$ is a forest. The *feedback edge number* $\mathrm{fen}(G)$ is the size of a smallest such set. Analogously, a set $V' \subseteq V(G)$ is a *feedback vertex set* if $G - V'$ is a forest. The *feedback vertex number* $\mathrm{fvn}(G)$ is the size of a smallest such set.

For a graph G, a *tree decomposition* is a pair (T, B), where T is a tree and $B\colon V(T) \to 2^{V(G)}$ such that (i) for each edge $uv \in E(G)$ there exists $x \in V(T)$ with $u, v \in B(x)$, and (ii) for each $v \in V(G)$ the set of nodes $x \in V(T)$ with $v \in B(x)$ forms a nonempty, connected subtree in T. The *width* of (T, B) is $\max_{x \in V(T)}(|B(x)| - 1)$. The *tree-width* $\mathrm{tw}(G)$ of G is the minimum width of all tree decompositions of G. The *path-width* $\mathrm{pw}(G)$ of G is the minimum width of all tree decompositions (T, B) of G for which T is a path.

The *tree-depth* of a connected graph G is defined as follows [22]. Let T be a rooted tree with vertex set $V(G)$, such that if $xy \in E(G)$, then x is either an ancestor or a descendant of y in T. We say that G is *embedded in* T. The *depth* of T is the number of vertices in a longest path in T from the root to a leaf. The *tree-depth* $\mathrm{td}(G)$ of G is the minimum t such that there is a rooted tree of depth t in which G is embedded.

We next define the *modular-width* of a graph G [16]. A vertex set $M \subseteq V(G)$ is a *module* if for all $v, w \in M$ it holds that $N(v) \cap (V(G) \setminus M) = N(w) \cap (V(G) \setminus M)$. We call a module M *trivial*, if $|M| \leq 1$ or $M = V(G)$, and we call it *strong* if for every other module M' of G we have that $M \cap M' = \emptyset$, or that one is a subset of the other. A graph that only admits trivial modules is called *prime*. Every non-singleton graph can be uniquely partitioned into maximal strong modules $\mathcal{P} = \{M_1, \ldots, M_\ell\}$ with $\ell \geq 2$. Recursively partitioning the graphs $G[M_i]$ in this way until every module is trivial yields a *modular decomposition* of G. The modular-width is the largest number of trivial modules in a *prime subgraph* $G[M_i]$ of the modular decomposition of G.

A *parameterized problem* is a subset $L \subseteq \Sigma^* \times \mathbb{N}$ over a finite alphabet Σ. Let $f\colon \mathbb{N} \to \mathbb{N}$ be a computable function. A problem L is *fixed-parameter tractable (in* FPT*)* with respect to k if $(I, k) \in L$ is decidable in time $f(k) \cdot |I|^{O(1)}$ and L is in XP if $(I, k) \in L$ is decidable in time $|I|^{f(k)}$. There is a hierarchy of computational complexity classes for parameterized problems: $\text{FPT} \subseteq \text{W}[1] \subseteq \text{W}[2] \subseteq \cdots \subseteq \text{XP}$. To show that a parameterized problem L is (presumably) not in FPT one may use a *parameterized reduction* from a W[i]-hard problem to L, for any $i \geq 1$. A parameterized reduction from a parameterized problem L to another parameterized problem L' is a function that acts as follows: For computable functions f and g, given an instance (I, k) of L, it computes in $f(k) \cdot |I|^{O(1)}$ time an instance (I', k') of L' so that $(I, k) \in L \iff (I', k') \in L'$ and $k' \leq g(k)$. For further information on fixed-parameter tractability and parameterized complexity we refer to Cygan et al. [11].

3 Hardness for Path-width and Feedback Vertex Number

In this section we show that Geodetic Set is W[1]-hard with respect to the feedback vertex number, the path-width and the solution size, combined. To this end, we present a parameterized reduction from Grid Tiling, which is W[1]-hard with respect to k [11, 21]:

Grid Tiling

Input: A collection $\mathcal{S}$ of k^2 sets $S^{i,j} \subseteq [m] \times [m]$, $i, j \in [k]$ (called tile sets), each of cardinality exactly n.

Question: Can one choose a tile $(x^{i,j}, y^{i,j}) \in S^{i,j}$ for each $i, j \in [k]$ such that $x^{i,j} = x^{i,j'}$ with $j' = (j + 1) \bmod k$ and $y^{i,j} = y^{i',j}$ with $i' = (i + 1) \bmod k$?

This distinguishes our reduction from most parameterized reductions to show W[1]-hardness, as one typically reduces from Clique, or its multicolored variant. Grid Tiling though seemed to be a much better fit, since the values of the tiles can be expressed by lengths of paths. This is the central idea for our reduction: We place a connection gadget between each pair of adjacent tile sets. Placing paths of fitting lengths, the connection gadget ensures that the vertices corresponding to the tiles agree with each other, that is, the appropriate coordinates of the two tiles are equal.

Remark. Throughout this section we write i' and j' as shorthands for $(i + 1) \bmod k$ and $(j + 1) \bmod k$, respectively. Moreover, we assume that the grid size k is even.

Construction. Let $I = (\mathcal{S}, k, m, n)$ be an instance of Grid Tiling. We construct an instance $I' = (G, k')$ of Geodetic Set as follows: First, we set $k' = k^2 + 4$. We add the *global vertices* $\Xi = \{\alpha, \beta, \gamma, \delta\}$ and $\Xi' = \{\alpha', \beta', \gamma', \delta'\}$, and add four edges $\alpha\alpha', \beta\beta', \gamma\gamma'$ and $\delta\delta'$. Next, for each $i, j \in [k]$ we introduce *tile vertices* $S^{i,j} = \{s_1^{i,j}, \ldots, s_n^{i,j}\}$. For a tile vertex v we denote by (x_v, y_v) the

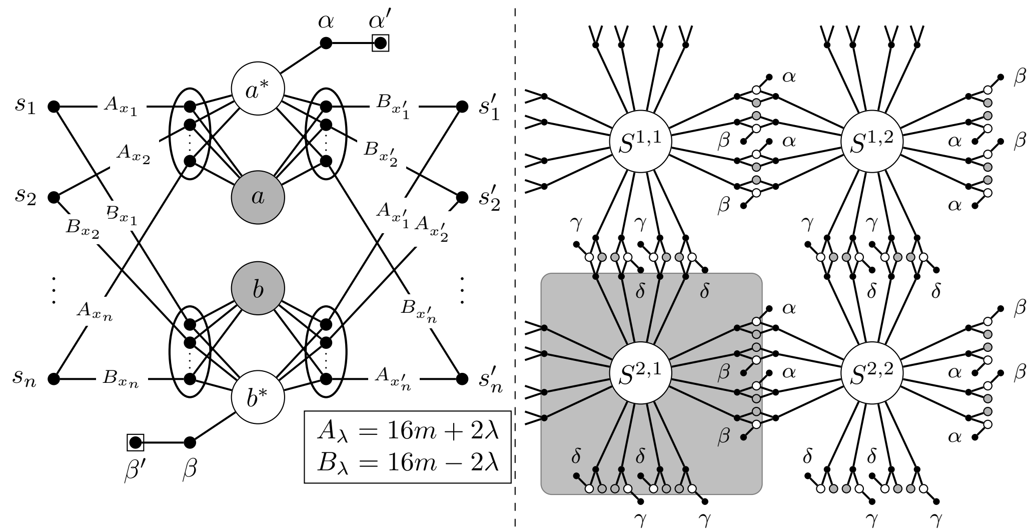

Figure 3: *Left:* One copy of a horizontal connection gadget next to $S^{i,j} = \{s_1, \dots, s_n\}$ where j is even, connecting the tile sets $S^{i,j}$ and $S^{i,j'}$. Edges with label ℓ in the figure represent paths of length ℓ. The ellipses mark the connector vertices towards $S^{i,j}$ and $S^{i,j'}$. *Right:* An exemplary reduction from an instance of Grid Tiling, where $k = 2$. Between every pair of horizontally, resp. vertically adjacent tile sets (big circles) there are two copies of horizontal, resp. vertical connection gadgets. Note that $\alpha, \beta, \gamma, \delta \in \Xi$ are global; every vertex labeled such is the same vertex. The gray square marks the vertices of $Q^{2,1}$ (note that $\beta, \delta \notin Q^{2,1}$). Note that this illustration wraps around its boundaries, that is, the edges on the left end connect to the vertices on the right end and the edges on the top end connect to the vertices on the bottom end.

corresponding tile. Moreover, for each $i, j \in [k]$ we introduce two copies of the horizontal and two copies of the vertical connection gadget.

The construction of a *horizontal connection gadget next to tile set* $S^{i,j}$ is as follows. Let $S = S^{i,j}$ and let $S' = S^{i,j'}$ be the vertices of the two horizontally adjacent tile sets. We introduce the vertices a and b called *hidden vertices* and the vertices a^* and b^* called *exposed vertices*. Next, for every tile vertex $s \in S$ with its corresponding tile (x_s, y_s), we add a path of length $16m + 2x_s + 1$ from s to a, and a path of length $16m - 2x_s + 1$ from s to b. For every tile vertex $s' \in S'$ with its corresponding tile $(x_{s'}, y_{s'})$, we add a path of length $16m - 2x_{s'} + 1$ from s' to a, and a path of length $16m + 2x_{s'} + 1$ from s' to b. We call these paths *tile paths towards* S, respectively S'. We call the neighbors of a, respectively b, *connector vertices towards* S, respectively S'. The exposed vertices a^*, respectively b^* are adjacent to all neighbors of a, respectively b. Moreover, each of a^* and b^* has one additional neighbor: If j is even, then α is a neighbor of a^* and β is a neighbor of b^*. If j is odd, then β is a neighbor of a^* and α is a neighbor of b^*. See Figure 3 (left) for an illustration of a horizontal connection gadget next to $S^{i,j}$ for even j.

The construction of a *vertical connection gadget next to tile set* $S^{i,j}$ is identical to the construction of a horizontal gadget, except for the following differences:

- the gadget connects tile sets $S = S^{i,j}$ and $S' = S^{i',j}$;
- the lengths of the tile paths depend on the y-coordinates; and

- if i is even, then γ is a neighbor of a^* and δ is a neighbor of b^*, and if i is odd, then δ is a neighbor of a^* and γ is a neighbor of b^*.

This concludes the construction. See Figure 3 (right) for an overview.

Let J be the set of all hidden vertices and let J^* be the set of all exposed vertices. We now show that this construction has the desired properties for showing W[1]-hardness with respect to solution size, feedback vertex number and path-width, combined.

Observation 1 *The constructed graph G has* $\mathrm{pw}(G) \leq 16k^2 + 2$ *and* $\mathrm{fvn}(G) \leq 16k^2$*.*

Proof: The graph $G' = G - (J \cup J^*)$ consists of paths of length one and subdivisions of stars. Clearly, $\mathrm{fvn}(G') = 0$, and since removing the center vertex of a subdivision of a star yields disjoint paths, $\mathrm{pw}(G') = 2$. Adding a vertex to a graph increases each of the two parameters by at most one. Now, as $|J \cup J^*| = 16k^2$, the claim follows. □

Correctness. Let us first point out that the computational challenge of the constructed Geodetic Set instance lies in finding vertices to cover all hidden vertices J, as every other vertex is covered by the four degree-one vertices in Ξ', which have to be in every solution as they cannot be covered in any other way.

Observation 2 $I[\Xi'] = V(G) \setminus J$*.*

Proof: For $i, j \in [k]$ and for $s \in S^{i,j}$ let $(x_s, y_s) \in [m] \times [m]$ be the values of the corresponding tile. We show first that all vertices in horizontal connection gadgets are covered. Suppose that j is even. For every $s \in S^{i,j'}$, there are 32 shortest paths of length $3 + 16m + 2x_s + 16m - 2x_s + 3 = 32m + 6$, each of which is also a shortest s-visiting path. Sixteen of the paths use horizontal connection gadgets, and sixteen paths use the vertical connection gadgets. Let us list the paths using the horizontal connection gadgets first. Denote by $X^{i,j}$, respectively $X^{i,j'}$, one of the two horizontal connection gadgets next to $S^{i,j}$, respectively $S^{i,j'}$. Let a^* and b^*, respectively a'^* and b'^*, be the two exposed vertices of $X^{i,j}$, respectively $X^{i,j'}$. Note that since j' is odd, a'^* is adjacent to β, while b'^* is adjacent to α. We find the following shortest α'–β'-paths via s and the two horizontal connection gadgets $X^{i,j}$ and $X^{i,j'}$: (1) one path via a^*, s, and b^*, (2) one path via b'^*, s, and a'^*, (3) one path via a^*, s, and a'^*, and (4) one path via b'^*, s, and b^*. Taking the copies of $X^{i,j}$ and $X^{i,j'}$, we find twelve further paths. Hence, overall there are sixteen shortest s-visiting α'–β'-paths that use horizontal connection gadgets.

The case that j is odd behaves analogously; note that α now is adjacent to the exposed vertex b^* while β is connected to a^*. Combining the two cases we conclude that the shortest α'–β'-paths cover all tile vertices as well as all vertices in horizontal connection gadgets, except for the hidden vertices.

By symmetry the shortest γ'–δ'-paths cover all tile vertices as well as all vertices in vertical connection gadgets, except for the hidden vertices; thus $V(G) \setminus J \subseteq I[\Xi']$.

It remains to be shown that $J \cap I[\Xi'] = \emptyset$. Note that the neighborhood of any hidden vertex is a subset of the neighborhood of the corresponding exposed vertex. Since each vertex in Ξ is adjacent to exactly one vertex in Ξ' and to exposed vertices, $I[\Xi']$ cannot contain any hidden vertex. □

Then the forward direction becomes straightforward: Our geodetic set V' consists of Ξ' and, for every tile in the solution of instance I, the corresponding tile vertex. It is easy to see that for every (copy of a) connection gadget, there are two shortest paths between the chosen tile vertices of any two adjacent tiles, each covering one of the two hidden vertices in the connection gadget. Compare with Figure 3 (hidden vertices are gray). We further derive the following observation, which is also the reason why the vertices in J are called *hidden*.

Observation 3 *Let $u, v \in V(G) \setminus (\Xi \cup \Xi')$. If a shortest u–v-path visits a global vertex, then none of its inner vertices are hidden.*

The backward direction is more involved. We show that every solution of our constructed instance consists of Ξ' and exactly one tile vertex of each tile set. For this we make use of two properties of our construction. First, if two vertices are sufficiently far apart, then there is a shortest path via some global vertex that connects them.

Lemma 1 *For any two vertices $u, v \in V(G)$ there is a u–v-path of length at most $36m + 6$ that visits some global vertex.*

Proof: We define $\xi_u \in \Xi$ as follows. If $u \in J \cup J^* \cup \Xi \cup \Xi'$, then let $\xi_u \in \Xi$ be an arbitrary global vertex such that $d(u, \xi_u) \leq d(u, \zeta)$ for all $\zeta \in \Xi$. Suppose that u is in a (horizontal or vertical) connection gadget. Then u lies on a path between a tile vertex $u' \in S^{i,j}$, and a connector vertex u'' towards $S^{i,j}$, where $i, j \in [k]$. Let $\xi_u \in \Xi$ be a global vertex such that $d(u'', \xi_u) \leq d(u'', \zeta)$, for $\zeta \in \Xi$. We define ξ_v analogously. If $\xi_u = \xi_v$, then $d(u, v) \leq d(u, \xi_u) + d(\xi_u, v) \leq 16m + 2\lambda + 2 + 2 + 2\lambda' + 16m \leq 36m + 6$, where $\lambda, \lambda' \in [m]$ are either x- or y-values of some tile.

So suppose that $\xi_u \neq \xi_v$. We will prove that

$$\begin{aligned} d(u, \xi_u) + d(\xi_u, v) + d(u, \xi_v) + d(\xi_v, v) &= d(\xi_u, u) + d(u, \xi_v) + d(\xi_v, v) + d(v, \xi_u) \\ &\leq 2(36m + 6), \end{aligned}$$

which yields the statement above as $d(u, v) \leq \min\{d(u, \xi_u) + d(\xi_u, v), d(u, \xi_v) + d(\xi_v, v)\}$. In particular, we show that $d(\xi_u, u) + d(u, \xi_v) \leq 36m + 6$. If $u \notin J$, then $d(\xi_u, u) + d(u, \xi_v) \leq d(\xi_u, u') + d(u', \xi_v)$ for some tile vertex u'. Thus we obtain

$$d(\xi_u, u) + d(u, \xi_v) \leq d(\xi_u, u') + d(u', \xi_v) = 2 + 16m + 2\lambda + 16m + 2\lambda' + 2 \leq 36m + 4,$$

where $\lambda, \lambda' \in [m]$ are either x- or y-values of some tile. If $u \in J$, then we have

$$d(\xi_u, u) + d(u, \xi_v) = 3 + 1 + 16m + 2\lambda + 16m + 2\lambda' + 2 \leq 36m + 6.$$

Analogously, $d(\xi_v, v) + d(v, \xi_u) \leq 36m + 6$, concluding the proof. □

We introduce some additional notation. The *square* $Q^{i,j}$ of tile set $S^{i,j}$ is the vertex set consisting of the tile vertices $S^{i,j}$, the paths between tile vertices and connector vertices towards $S^{i,j}$, and all hidden vertices and exposed vertices that are in the connection gadgets next to $S^{i,j}$. See Figure 3 (right) for an illustration of a square. Note that the squares are pairwise disjoint. We say that two squares are adjacent if they contain vertices of the same connection gadget. The *adjacency* $\mathrm{Adj}(Q^{i,j})$ of a square $Q^{i,j}$ is the union of squares adjacent to $Q^{i,j}$. The *closed adjacency* of a square $Q^{i,j}$ is the vertex set $\mathrm{Adj}[Q^{i,j}] = \mathrm{Adj}(Q^{i,j}) \cup Q^{i,j}$.

We will show that any solution of (G, k') contains exactly one vertex per square. Our proof has two parts. Assume that there is a solution V' such that there is a square that contains no vertex from V'. We call such a square *empty*. In the first part (Lemma 2), we show that, if there is at least one empty square, then for one of them, the adjacent squares contain at most eight vertices from V'. We then argue that this contradicts the fact that V' is geodetic in the second part (Lemma 3).

The following lemma pertains to the first part. We remark that this may be of independent interest, as this may turn out useful when proving the correctness of a reduction from Grid Tiling. For the lemma we use the following notation: For a $k \times k$ matrix M with entries $m_{i,j}$, $i, j \in [k]$ let $\delta^M_{i,j}$ be the sum of the entries that are above, below, to the left, and to the right of $m_{i,j}$, that is, $\delta^M_{i,j} = m_{i',j} + m_{i'',j} + m_{i,j'} + m_{i,j''}$, where $i' = (i+1) \bmod k$, $i'' = (i-1) \bmod k$, $j' = (j+1) \bmod k$, and $j'' = (j-1) \bmod k$.

Lemma 2 *Let $A \in \mathbb{N}^{k\times k}$ be a matrix with even k, such that $\sum_{i,j\in[k]} a_{i,j} = k^2$. Then, there exist $i, j \in [k]$ such that $a_{i,j} = 0$ and $\delta^A_{i,j} \leq 8$, unless $a_{i,j} = 1$ for all $i, j \in [k]$.*

Proof: Let $q > 0$ be the number of zero entries in A. We show that $\sum_{i,j\in[k],a_{i,j}=0} \delta^A_{i,j} \leq 8q$. The lemma then follows by the pigeonhole principle.

Let $B, C \in \mathbb{N}^{k\times k}$ be matrices such that for every $i, j \in [k]$,

$$b_{i,j} = \begin{cases} 1 & \text{if } a_{i,j} \geq 1 \\ 0 & \text{otherwise,} \end{cases} \qquad c_{i,j} = \begin{cases} a_{i,j} - 1 & \text{if } a_{i,j} \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Observe that $A + B = C$ and thus $\delta^A_{i,j} = \delta^B_{i,j} + \delta^C_{i,j}$ for every $i, j \in [k]$. We will prove that (1) $\sum_{i,j\in[k],a_{i,j}=0} \delta^B_{i,j} \leq 4q$, and (2) $\sum_{i,j\in[k],a_{i,j}=0} \delta^C_{i,j} \leq 4q$. Note that $\delta^B_{i,j} \leq 4$ for every $i, j \in [k]$, as $b_{i,j} \leq 1$. This proves (1). For (2), note that $\sum_{i,j\in[k]} c_{i,j} = \sum_{i,j\in[k],a_{i,j}\geq 1}(a_{i,j}-1) = k^2-(k^2-q) = q$ since $\sum_{i,j\in[k]} a_{i,j} = k^2$ and A has k^2-q non-zero entries. Since $\sum_{i,j\in[k],a_{i,j}=0} \delta^C_{i,j} = \sum_{i,j\in[k]} z_{i,j}c_{i,j}$ where $z_{i,j} \in \{0, 1, \ldots, 4\}$ for every $i, j \in [k]$ (every $c_{i,j}$ is adjacent to $z_{i,j} \leq 4$ zero entries), we have $\sum_{i,j\in[k],a_{i,j}=0} \delta^C_{i,j} \leq 4 \sum_{i,j\in[k]} c_{i,j} \leq 4q$. □

We remark that the upper bound is tight: Consider a matrix with 0 and 2 entries arranged in a chessboard layout.

Lemma 3 *A geodetic set $V' \subseteq V(G)$ of size at most k' consists of the four vertices in Ξ', and exactly one vertex in each square $Q^{i,j}$, for each $i, j \in [k]$.*

Proof: Recall that $k' = k^2 + 4$. The four vertices in Ξ' are the only vertices of degree one and are part of every geodetic set. Further we may assume that $V' \cap \Xi = \emptyset$ as $I[V'] = I[V' \setminus \Xi]$. So V' consists of the four vertices in Ξ' and a set of at most k^2 vertices within the squares, denoted by W.

For contradiction, assume that there is a square Q' which is *empty* (that is, $Q' \cap W = \emptyset$). Then, by Lemma 2, there exists an empty square Q for which $|\operatorname{Adj}(Q) \cap W| \leq 8$.

Let $J_Q = J \cap N[Q]$ be the sixteen hidden vertices that are either in Q or adjacent to vertices of Q. The next two claims are consequences of Lemma 1 and Observation 3:

(1) no shortest path between a vertex outside of Q and a vertex outside of $\operatorname{Adj}[Q]$ can visit any vertex in J_Q, and

(2) W covers at most $|\operatorname{Adj}(Q) \cap W| \leq 8$ vertices of J_Q.

For (1), let $u \in V(G) \setminus Q$, let $v \in J_Q$, and let $w \in V(G) \setminus \operatorname{Adj}[Q]$ (possibly $v = w$). Observe first that any shortest path that visits v and whose endpoints are not in the connection gadget containing v visits tile vertices of both incident tiles. Also, the shortest path cannot visit any of the global vertices Ξ as they provide a short cut around v. Then it is easy to see that any shortest u–w-path visiting v must at some point go through vertices of some square $Q' \subseteq \operatorname{Adj}(Q)$, then visit v, enter Q, and then leave Q again before reaching w. Within Q', such a path covers a distance of at least $16m - 2\lambda + 16m - 2\lambda' \geq 28m$ for $\lambda, \lambda' \in [m]$. Analogously, it covers a distance of at least $28m$ within Q as well. Thus its length is at least $56m$, and by Lemma 1 such a path is longer than $\operatorname{diam}(G)$, contradicting the existence of a shortest u-w-path that visits a vertex in J_Q.

For (2), suppose that there exist $u, u' \in \operatorname{Adj}(Q)$ such that there is a shortest u–u'-path P that visits $v \in J_Q$. It is easy to see that P must go through $v' \neq v \in J_Q$. Assume without loss of generality that v appears before v' in P. In order for P to be a shortest path, it must hold that

$$d(u, v) + d(v, v') + d(v', u') \leq d(u, u') \leq 36m + 6,$$

due to Lemma 1. Since $d(v, v') = 2 + 16m - 2\lambda + 16m - 2\lambda' + 2$ for some $\lambda, \lambda' \in [m]$, we have $d(v, v') \geq 28m+4$; thus we can assume that $d(u, v)+d(u', v') \leq 8m+2$. Then, by construction, u (respectively u') lies on a path between a tile vertex and v (respectively v'). By (1), only the vertices in $W \cap \operatorname{Adj}(Q)$ can cover the vertices in J_Q. Hence, for every vertex $u \in W \cap \operatorname{Adj}(Q)$ there is at most one vertex $v \in J_Q$ that is going to be in $I[W]$, and the claimed inequality holds.

Since $|J_Q| = 16$, the set V' is not geodetic; so there cannot be an empty square in G. There are k^2 squares and $|W| = |V' \setminus \Xi'| \leq k^2$. So $|V' \cap Q^{i,j}| = 1$ for each $i, j \in [k]$. □

Using Lemma 3, we show that every solution vertex in a square must be a tile vertex.

Lemma 4 *A geodetic set $V' \subseteq V(G)$ of size at most k' consists of the four vertices in Ξ' and exactly one vertex of $S^{i,j}$, for each $i, j \in [k]$.*

Proof: For $i, j \in [k]$, let $S = S^{i,j}$, $S' = S^{i,j'}$, $Q = Q^{i,j}$, and $Q' = Q^{i,j'}$. Without loss of generality, assume that j is even (see Figure 3 for an illustration). Let X_1 and X_2 be the two copies of the horizontal connection gadget next to tile S, let $a_1, b_1 \in V(X_1)$ and $a_2, b_2 \in V(X_2)$ be the hidden vertices, and let $a_1^*, b_1^* \in V(X_1)$ and $a_2^*, b_2^* \in V(X_2)$ be the exposed vertices. By Lemma 3, V' contains exactly one vertex u in Q and exactly one vertex v in Q'. Let us fix these vertices for the remainder of the proof.

Consider a vertex $w \in V(G) \setminus (Q \cup Q')$. Note that any shortest u–w-path and any shortest v–w-path going through one of a_1, a_2, b_1, b_2 must use tile vertices in S and S'. It is easy to verify that due to its length, such a path must visit some global vertex, thus it cannot visit any hidden vertex (Observation 3). It follows that $\{a_1, a_2, b_1, b_2\} \subseteq I[u, v]$.

For the sake of contradiction, suppose that $u \notin S$. In particular, we assume without loss of generality that $u \in V(X_1)$. Let $u' \in S$ be the tile vertex such that u lies on the tile path between u' and a_1. Observe that $d(u, a_1) < d(u, a_2)$. Hence, no shortest u–v-path visits a_2 if $d(v, a_1) \leq d(v, a_2)$. It follows that v lies on some tile path between some tile vertex $v' \in S'$ and a_2. Since there are shortest u–v-paths visiting a_1 and a_2, we have

$$\begin{aligned} d(u, v) &= (d(a_1, u') - d(u', u)) + d(a_1, v') + d(v, v') \text{ and} \\ d(u, v) &= (d(a_2, v') - d(v, v')) + d(a_2, u') + d(u, u'). \end{aligned}$$

By construction, $d(a_1, u') = d(a_2, u') = 16m + 2x_{u'} + 1$ and $d(a_1, v') = d(a_2, v') = 16m - 2x_{v'} + 1$. Thus, we obtain $d(u, u') = d(v, v')$ and $d(u, v) = 32m + 2x_{u'} - 2x_{v'} + 2$. Note that there is a u–v-path visiting α that is of length

$$\ell = (d(u', a_1^*) - d(u, u')) + 2 + (d(a_2^*, v') - d(v', v)).$$

Since $d(a_1, u') = 16m + 2x_{u'} + 1$ and $d(v', a_1) = 16m - 2x_{v'} + 1$ (by construction), and since $\ell \geq d(u, v)$, we obtain $d(u, u') = d(v, v') \leq 1$. By the assumption that $u \notin S$, we have $d(u, u') > 0$.

It follows that $d(u, u') = d(v, v') = 1$. Finally, observe that the shortest path from u to v that visits b_1 is of length

$$\ell' = d(u, b_1) + d(b_1, v) = 32m - 2x_{u'} + 2x_{v'} + 4.$$

Since $\ell' = d(u, v)$, we obtain $4x_{u'} - 4x_{v'} = 2$, so one of $x_{u'}, x_{v'}$ cannot be integer; a contradiction. □

Now, given Lemma 4, if there is a solution for our instance of Geodetic Set, then the tiles corresponding to the chosen tile vertices are a solution for our instance of Grid Tiling. The main theorem of the section follows:

Theorem 4 Geodetic Set *is* W[1]*-hard with respect to the feedback vertex number, the path-width, and the solution size, combined.*

Proof: Given an instance $(\mathcal{S}, k, m, n)$ of Grid Tiling, we construct an instance (G, k') of Geodetic Set as shown above. We now prove that $(\mathcal{S}, k, m, n)$ is a yes instance if and only if (G, k') is a yes-instance.

For the *only if* direction, let $\mathcal{S}' = \{(x^{i,j}, y^{i,j}) \in S^{i,j} \mid i, j \in [k]\}$ be a solution for the instance $(\mathcal{S}, k, m, n)$. Then we construct a geodetic set V' by adding the vertices in Ξ' and, for every $i, j \in [k]$, the tile vertex $s^{i,j} \in S^{i,j}$, corresponding to $(x^{i,j}, y^{i,j})$. Clearly, $|V'| = k' = k^2 + 4$. By Observation 2, all vertices in $V(G) \setminus J$ are covered. For $i, j \in [k]$, let a and b be the hidden vertices of one of the copies of the horizontal connection gadget next to $S^{i,j}$. Since $x^{i,j} = x^{i,j'}$, the shortest a-visiting $s^{i,j}$–$s^{i,j'}$-paths have length

$$d(s^{i,j}, a, s^{i,j'}) = 16m + 2x^{i,j} + 1 + 1 + 16m - 2x^{i,j'} = 32m + 2,$$

and the shortest b-visiting $s^{i,j}$–$s^{i,j'}$-paths have length

$$d(s^{i,j}, b, s^{i,j'}) = 16m - 2x^{i,j} + 1 + 1 + 16m + 2x^{i,j'} = 32m + 2.$$

It is easy to see that there are no shorter $s^{i,j}$–$s^{i,j'}$-paths. So the hidden vertices of the two copies of the horizontal connection gadget are in $I[V']$. Analogously, since $y^{i,j} = y^{i',j}$, there exist shortest $v^{i,j}$–$v^{i',j}$-paths that visit the hidden vertices of the two copies of the vertical connection gadgets next to $S^{i,j}$. Thus V' is geodetic and of cardinality k'.

For the *if* direction, let V' be a solution for (G, k'). By Lemma 4, V' consists only of the vertices in Ξ' and one tile vertex $s^{i,j}$ for each $i, j \in [k]$. Let $(x^{i,j}, y^{i,j})$ be the corresponding pair. Note that the hidden vertices within the copies of $X^{i,j}$ can only be covered by shortest $s^{i,j}$–$s^{i,j'}$-paths. But in order for these paths to be of equal length, it must hold that $x^{i,j} = x^{i,j'}$. Analogously, in order to cover the hidden vertices within the copies of $Y^{i,j}$, we must have $y^{i,j} = y^{i',j}$. So choosing the pair $(x^{i,j}, y^{i,j})$ for each $i, j \in [k]$ yields a solution for the instance $(\mathcal{S}, k, m, n)$ of Grid Tiling.

Since Grid Tiling is W[1]-hard with respect to k, it follows from the reduction and Observation 1 that Geodetic Set is W[1]-hard with respect to $k' + \text{fvn}(G) + \text{pw}(G)$. □

4 Fixed-Parameter Tractability for Feedback Edge Number

We now show that Geodetic Set is fixed-parameter tractable for feedback edge number. In fact, we present a fixed-parameter algorithm for the following, more general variant:

Extended Geodetic Set

Input: A graph G, a vertex set $T \subseteq V(G)$, and an integer k.
Question: Does G have a geodetic set $S \supseteq T$ of cardinality at most k?

The algorithm works in three steps: We first apply some polynomial-time data reduction rules (Section 4.1), after which the graph may still be arbitrarily large but is easier to handle due to its structure. Afterwards, with some branching steps (Section 4.2), we obtain an instance in which a part of the solution vertices are fixed and can be extended to a minimum geodetic set by adding vertices on paths of degree-two vertices. We determine where on the paths to place these vertices using an ILP formulation with $O(\text{fen}(G)^2)$ variables (Section 4.3), showing that (Extended) Geodetic Set is fixed-parameter tractable for feedback edge number.

Although feedback edge number is considered one of the largest structural graph parameters, our algorithm is still technically involved and it has an impractical running time. This hints at the difficulty of designing efficient algorithms for Geodetic Set. We also remark that some of the techniques presented may be of independent interest. For example, the presented approach may also be useful to show fixed-parameter tractability of the closely related Metric Dimension problem[2] for feedback edge number, which was posed as an open problem by Eppstein [14] (so far, it is only known to be in XP for this parameter [15]).

Throughout this section we assume without loss of generality that G is connected.

4.1 Preprocessing

We next present three data reduction rules and some observations on the instance obtained after their exhaustive application. We will also introduce the *feedback edge graph* $\widetilde{G}$ in this subsection, which will be used throughout the presentation of this algorithm.

Our first reduction rule deletes degree-one vertices. This reduction rule is based on the observation that a geodetic set contains every degree-one vertex.

Reduction Rule 1 *If there is a degree-one vertex $v \in V(G)$ with $N(v) = \{u\}$, then*

- *decrease k by 1 if $u \in T$,*
- *add u to T if $u \notin T$, and*
- *delete v from $V(G)$ (and from T).*

Henceforth we assume that Reduction Rule 1 has been exhaustively applied (which can be done in linear time). Suppose that $\text{fen}(G) = 1$. Then G is a cycle, and any minimal geodetic set $S \supseteq T$ is of size at most $|T| + 3$. So Extended Geodetic Set can be solved in polynomial time when $\text{fen}(G) \leq 1$ (in fact, further analysis yields a linear-time algorithm for $\text{fen}(G) = 1$). We thus assume that $\text{fen}(G) \geq 2$.

Now we introduce the *feedback edge graph* $\widetilde{G}$, a multigraph which is obtained from G as follows: As long as there is a degree-two vertex v with neighbors u, w, we remove v and add an edge (multiedge) uw. Using the handshake lemma, one can easily obtain the following.

Observation 5 *It holds that $|V(\widetilde{G})| \leq 2\,\text{fen}(G) - 2$ and $|E(\widetilde{G})| \leq 3\,\text{fen}(G) - 3$.*

Proof: By definition, $|E(G)| \leq |V(G)| + \text{fen}(G) - 1$. It follows that $|E(\widetilde{G})| \leq |V(\widetilde{G})| + \text{fen}(G) - 1$, since the number of edges decreases by 1 every time we remove a vertex. By the handshake lemma, $2|E(\widetilde{G})| = \sum_{v \in V(\widetilde{G})} \deg_{\widetilde{G}}(v) \geq 3|V(\widetilde{G})|$. Solving the inequalities for $|V(\widetilde{G})|$ and $|E(\widetilde{G})|$ respectively yields the sought bounds. □

[2] Given a graph, Metric Dimension asks for a set S of at most k vertices such that for any pair of vertices u and v, there is a vertex in S which has distinct distances to u and v.

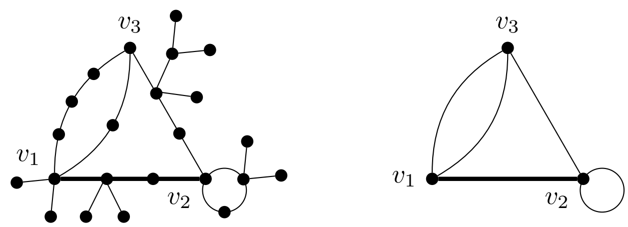

Figure 4: An illustration of an input graph G (left) and $\widetilde{G}$ after Reduction Rule 1 has been exhaustively applied (right). Observe that $\widetilde{G}$ contains no degree-one or degree-two vertex. For instance, a thick edge p in $\widetilde{G}$ (right) corresponds to a path P of length $h_p = 3$ in G(left). Moreover, we have $T_p = \{0, 1\}$ after Reduction Rule 1 has been applied exhaustively.

Observe that each edge p in $\widetilde{G}$ is associated with a path $P = (p^0, p^1, \dots, p^{h_p})$ in G where all of its inner vertices are of degree 2. We sometimes refer to the endpoints p^0 and p^{h_p} as $p^{\leftarrow}$ and $p^{\rightarrow}$, respectively. Moreover, let $T_p = \{i \mid p^i \in T\}$ and let $p_T^{\leftarrow} = p^{t_p^{\leftarrow}}$ and $p_T^{\rightarrow} = p^{t_p^{\rightarrow}}$, where $t_p^{\leftarrow} = \min T_p$ and $t_p^{\rightarrow} = \max T_p$. We illustrate the definitions in Figure 4.

The following reduction rule deals with self-loops in $\widetilde{G}$.

Reduction Rule 2 *If $v \in V(\widetilde{G})$ has a self-loop p in $\widetilde{G}$, then decrease k as follows:*

- *If $T_p = \emptyset$, then decrease k by $(h_p \bmod 2)$.*
- *If $T_p \neq \emptyset$ and $V(P) \not\subseteq I[T_p \cup \{v\}]$, then decrease k by $|T_p|$.*
- *If $T_p \neq \emptyset$ and $V(P) \subseteq I[T_p \cup \{v\}]$, then decrease k by $|T_p| - 1$.*

Moreover, add v to T and remove $V(P) \setminus \{v\}$.

Lemma 5 *Reduction Rule 2 is correct.*

Proof: We reduce the first two cases to the third case with the following observations:

- If $T_p = \emptyset$, then (G, T, k) is equivalent to $(G, T' = T \cup \{p^{\lfloor h_p/2 \rfloor}, p^{\lceil h_p/2 \rceil}\}, k)$. Then $V(P) \subseteq I[T'_p \cup \{v\}]$ and $|T'_p| - 1 = (h_p \bmod 2)$.
- If $T_p \neq \emptyset$ and $V(P) \not\subseteq I[T_p \cup \{v\}]$, then it is equivalent either to $(G, T' = T \cup \{p^{\lfloor h_p/2 \rfloor}\}, k)$ or to $(G, T'' = T \cup \{p^{\lceil h_p/2 \rceil}\}, k)$. Then $V(P) \subseteq I[T'_p \cup \{v\}]$ or $V(P) \subseteq I[T''_p \cup \{v\}]$ and $|T'_p| - 1 = |T''_p| - 1 = |T_p|$.

So assume that $T_p \neq \emptyset$ and $V(P) \subseteq I[T_p \cup \{v\}]$.

Let (G', T', k') be an Extended Geodetic Set instance as a result of Reduction Rule 2. Note that $G' = G - (V(P) \setminus \{v\})$, $T' = T \cup \{v\}$, and $k' = k - |T_p| + 1$. It is easy to see that if $S \supseteq T$ is geodetic in G and $|S| \leq k$, then $(S \setminus V(P)) \cup \{v\}$ is a solution of (G', T', k'). Conversely, if $S' \supseteq T'$ is a geodetic set in G' of size at most k', then $(S' \setminus \{v\}) \cup T_p$ is a solution of (G, T, k). □

The next reduction rule ensures that for every $p \in E(\widetilde{G})$ with $T_p \neq \emptyset$, there is a shortest path from an endpoint of P to the closest vertex in T_p that is contained inside P. For this we introduce the following notation. Let $\mathcal{R} = \{\leftarrow, \rightarrow\}$. For $r \in \mathcal{R}$, we denote by $\bar{r} \in \mathcal{R} \setminus \{r\}$ the opposite direction.

Reduction Rule 3 *Let $p \in E(\widetilde{G})$ with $T_p \neq \emptyset$, and let $r \in \mathcal{R}$. If $d_P(p_T^r, p^r) > d_P(p_T^r, p^{\overline{r}}) + d_G(p^{\overline{r}}, p^r)$, then add p^q to T, where p^q is between p_T^r and p^r and $d(p^q, p_T^r) = \lfloor (h_p + d_G(p^{\leftarrow}, p^{\rightarrow}))/2 \rfloor$.*

Lemma 6 *Reduction Rule 3 is correct.*

Proof: Suppose that (G, T, k) is a yes-instance with a solution $S \supseteq T$. Let P' be a subpath of P with endpoints p^q and p_T^r. Note that

$$\begin{aligned} d_P(p^q, p^r) + d_G(p^r, p^{\overline{r}}) + d_P(p^{\overline{r}}, p_T^r) &= (t_P^r - d_P(p^q, p_T^r)) + d_G(p^r, p^{\overline{r}}) + (h_p - t_P^r) \\ &= h_p + d_G(p^r, p^{\overline{r}}) - d(p^q, p_T^r) \\ &= \lceil (h_p + d_G(p^r, p^{\overline{r}}))/2 \rceil \geq d_P(p^q, p_T^r). \end{aligned}$$

Thus, S must contain a vertex $v \in V(P) \setminus \{p_T^r\}$ to cover P'. The correctness follows, because $(S \setminus \{v\}) \cup \{p^q\}$ is also geodetic in G. □

4.2 Guessing

We next extend our current set T of vertices fixed in the solution. First we guess the set of path endpoints that are in the solution. Next, using another reduction rule, we fix further vertices that are required to be in the geodetic set of our interest. These vertices possibly depend on the (previously guessed) endpoints that are in the solution. Finally, we guess how many vertices we need to add to every path P for $p \in E(\widetilde{G})$. Then, the exact positions of these vertices are determined using ILP.

Suppose that (G, T, k) is a yes-instance. We fix a solution S of minimum size that maximizes the number $|S \cap V(\widetilde{G})|$ of endpoints among all such solutions. Intuitively, our goal is to find S. To do so, we first guess the set $\widetilde{S} = S \cap V(\widetilde{G})$ of endpoints in S; there are at most $2^{|V(\widetilde{G})|} \leq 2^{2\,\mathrm{fen}(G)-2}$ possibilities by Observation 5. We extend T by adding all vertices from $\widetilde{S}$. So we will henceforth assume that $S \cap V(\widetilde{G}) = T \cap V(\widetilde{G})$. Using another reduction rule, we ensure that for every $p \in E(\widetilde{G})$, the vertices between $p_T^{\leftarrow}$ and $p_T^{\rightarrow}$ are covered.

Reduction Rule 4 *Let $p \in E(\widetilde{G})$. If there are $t < t' \in T_p$ such that $[t+1, t'-1] \cap T_p = \emptyset$ and $d_G(p^t, p^{t'}) < t' - t$ (equivalently, $d_G(p^{\leftarrow}, p^{\rightarrow}) + h_p < 2t' - 2t$), then add $p^{\lfloor (t+t')/2 \rfloor}$ to T.*

Lemma 7 *Reduction Rule 4 is correct.*

Proof: Let S be a geodetic set and let $V_p = \{p^{t+1}, \ldots, p^{t'-1}\}$. It suffices to show that $S \cap V_p \neq \emptyset$ and that $S' = (S \setminus V_p) \cup \{p^{\lfloor (t+t')/2 \rfloor}\}$ is geodetic. Suppose that $S \cap V_p = \emptyset$. Then for each $v, v' \in S$, no shortest path between v and v' visits a vertex in V_i. Hence, we have $S \cap V_p \neq \emptyset$. For the latter part, it is easy to see that S' is geodetic because $V_p \subseteq I[\{p^t, p^{\lfloor (t+t')/2 \rfloor}, p^{t'}\}]$. □

We will prove two lemmata required for the next guessing step and for the subsequent ILP formulation. First, we show that S contains no vertex on a path P for $p \in E(\widetilde{G})$ with $T_p \neq \emptyset$.

Lemma 8 *Let $p \in E(\widetilde{G})$ with $T_p \neq \emptyset$. Then, $S \cap V(P) \subseteq T_p$.*

Proof: For $r \in \mathcal{R}$, suppose that S contains a vertex $p^i \in V(P) \setminus T_p$ that lies between p^r and p_T^r. Since Reduction Rule 3 is applied exhaustively, $(S \setminus \{p^i\}) \cup \{p^r\}$ is also a solution of minimum size, contradicting the maximality of $|S \cap V(\widetilde{G})|$. Thus, it remains to show that S contains no vertex that lies between $p_T^{\leftarrow}$ and $p_T^{\rightarrow}$ in P. Note that after applying Reduction Rule 4, each vertex in P between $p_T^{\leftarrow}$ and $p_T^{\rightarrow}$ are included in $I[T_p]$. Due to its minimality, S contains no vertex $p^i \in V(P) \setminus T_p$ between $p_T^{\leftarrow}$ and $p_T^{\rightarrow}$ in P. □

We also show that S contains at most two inner vertices of P if $T_p = \emptyset$ for $p \in E(\widetilde{G})$.

Lemma 9 *Let $p \in E(\widetilde{G})$ with $T_p = \emptyset$. Then, $|S \cap V(P)| \leq 2$.*

Proof: If $|S \cap V(P)| = 3$, then $(S \setminus V(P)) \cup \{p^{\leftarrow}, p^{\lfloor h_p/2 \rfloor}, p^{\rightarrow}\}$ is also a minimum solution, contradicting the fact that $|S \cap V(\widetilde{G})|$ is maximized. □

Now we make further guesses. For each edge $p \in E(\widetilde{G})$, we guess the number $n_p \in \{0, 1, 2\}$ of inner vertices in $S \cap V(P)$. Note that there are at most $3^{|E(\widetilde{G})|} \leq 3^{3\,\mathrm{fen}(G)-3}$ possibilities by Observation 5. The next step is to determine exactly which vertices to take using ILP.

4.3 Finding a minimum geodetic set via ILP

Let $E_n = \{p \in E(\widetilde{G}) \mid T_p = \emptyset, n_p = n\}$ for $n \in \{0, 1, 2\}$ and let $E' = \{p \in E(\widetilde{G}) \mid T_p \neq \emptyset\}$. Further, let $\mathcal{E} = E_1 \cup E_2 \cup E' = E(\widetilde{G}) \setminus E_0$. Note that S contains at least one vertex in $V(P)$ for every $p \in \mathcal{E}$. For each $p \in \mathcal{E}$, we introduce two nonnegative variables $x_p^{\leftarrow}, x_p^{\rightarrow}$, and let $p_S^{\leftarrow} = p^{x_p^{\leftarrow}}$ and $p_S^{\rightarrow} = p^{h_p - x_p^{\rightarrow}}$. The intended meaning of $x_p^{\leftarrow}$, respectively $x_p^{\rightarrow}$ is that S contains $p_S^{\leftarrow}$, respectively $p_S^{\rightarrow}$. Then the geodetic set of our interest will be given by $X = T \cup \bigcup_{p \in E_1 \cup E_2} \{p_S^{\leftarrow}, p_S^{\rightarrow}\}$. For each $p \in \mathcal{E}$ we add the following basic constraints:

$$\begin{cases} x_p^{\leftarrow} > 0, x_p^{\rightarrow} > 0, \text{and } x_p^{\leftarrow} + x_p^{\rightarrow} \leq h_p & \text{if } p \in E_1 \cup E_2, \\ x_p^{\leftarrow} + x_p^{\rightarrow} = h_p & \text{if } p \in E_1, \\ h_p - 2x_p^{\leftarrow} - 2x_p^{\rightarrow} \leq d_G(v_p^{\leftarrow}, v_p^{\rightarrow}) & \text{if } p \in E_2, \\ x_p^{\leftarrow} = p_T^{\leftarrow} \text{ and } x_p^{\rightarrow} = h_p - p_T^{\rightarrow} & \text{if } p \in E'. \end{cases} \tag{1}$$

Let $V_p^{\leftarrow} = \{p^1, \ldots, p^{x_p^{\leftarrow}-1}\}$ and $V_p^{\rightarrow} = \{p^{h_p - x_i^{\rightarrow}+1}, \ldots, p^{h_p-1}\}$ for each $p \in \mathcal{E}$. We show that Constraint (1) guarantees that the vertices between $p_S^{\leftarrow}$ and $p_S^{\rightarrow}$ are covered if $p \notin E_0$.

Lemma 10 *If Constraint (1) is fulfilled, then $Q_p = V(P) \setminus (\{p^{\leftarrow}, p^{\rightarrow}\} \cup V_p^{\leftarrow} \cup V_p^{\rightarrow}) \subseteq I[S]$ holds for each $p \in \mathcal{E}$.*

Proof: If $p \in E_1$, then we have $Q_p = \{p^{x_p^{\leftarrow}}\} = \{p^{x_p^{\rightarrow}}\}$ and hence $Q_p \subseteq I[S]$. If $p \in E_2$, then we have $Q_p = \{p^{x_p^{\leftarrow}}, p^{x_p^{\leftarrow}+1}, \ldots, p^{x_p^{\rightarrow}}\}$. It follows from Constraint (1) that $d_P(p_S^{\leftarrow}, p_S^{\rightarrow}) \leq d_P(p_S^{\leftarrow}, p^{\leftarrow}) + d_G(p^{\leftarrow}, p^{\rightarrow}) + d_P(p^{\rightarrow}, p_S^{\rightarrow})$. This implies that $Q_p \subseteq I[p_S^{\leftarrow}, p_S^{\rightarrow}] \subseteq I[S]$. Finally, if $p \in E'$, then all vertices in Q_p are covered as shown in Lemma 8. □

It remains to work out the constraints to cover (i) $V(\widetilde{G}) \setminus T$, (ii) the inner vertices of P for $p \in E_0$, and (iii) $V_p^{\leftarrow} \cup V_p^{\rightarrow}$ for $p \in \mathcal{E}$. First, we describe linear constraints to check whether an endpoint of $V(\widetilde{G})$ is visited by a shortest path starting from its inner path vertex. The other endpoint of the path then is either in the same path (Lemma 12) or in a different path (Lemma 11).

Lemma 11 *Let $p \neq q \in E(\widetilde{G})$ and $r, s \in \mathcal{R}$ (recall that $\mathcal{R} = \{\leftarrow, \rightarrow\}$). Then, there is a shortest path between p_S^r and q_S^s visiting p^r and q^s if and only if the following hold:*

$$\begin{cases} x_p^r + d_G(p^r, q^s) + x_q^s \leq x_p^r + d_G(p^r, q^{\overline{s}}) + h_q - x_q^s, \\ x_p^r + d_G(p^r, q^s) + x_q^s \leq h_p - x_p^r + d_G(p^{\overline{r}}, q^s) + x_q^s, \\ x_p^r + d_G(p^r, q^s) + x_q^s \leq h_p - x_p^r + d_G(p^{\overline{r}}, q^{\overline{s}}) + h_q - x_q^s. \end{cases} \tag{2}$$

Proof: The length of the shortest p^r_S–q^s_S-path visiting p_r and q_r is $x^r_p + d_G(p^r, q_s) + x^s_q$. This path is shortest when it is (not necessarily strictly) shorter than p^r_S–q^s_S-paths visiting $p_{\overline{r}}$ or $q_{\overline{s}}$. □

Lemma 12 *Let $p \in E(\widetilde{G})$. Then, there is a shortest path between $p^{\leftarrow}_S$ and $p^{\rightarrow}_S$ visiting $p^{\leftarrow}$ and $p^{\rightarrow}$ if and only if the following holds:*

$$x^{\leftarrow}_p + d_G(p^{\leftarrow}, p^{\rightarrow}) + x^{\rightarrow}_p \leq h_p - x^{\leftarrow}_p - x^{\rightarrow}_p. \tag{3}$$

Proof: The length of the shortest $p^{\leftarrow}_S$–$p^{\rightarrow}_S$-path visiting $p^{\leftarrow}$ and $p^{\rightarrow}$ is $x^{\leftarrow}_p + d_G(p^r, q_s) + x^{\leftarrow}_q$. This path is shortest when it is (not necessarily strictly) shorter than any $p^{\leftarrow}_S$–$p^{\rightarrow}_S$-path visiting only inner vertices of P, whose length is $h_p - x^{\leftarrow}_p - x^{\rightarrow}_p$. □

To cover the remaining vertices, observe the following.

Observation 6 *The following hold.*

(i) For every $v \in V(\widetilde{G}) \setminus T$, v is covered if and only if there exist $(p, r) \neq (q, s) \in \mathcal{E} \times \mathcal{R}$ such that a shortest p^r–q^s-path visits v and a shortest p^r_S–q^s_S-path visits p^r and q^s (that is, $d(p^r, v) + d(v, q^s) = d(p^r, q^s)$).

(ii) For every $p \in E_0$ where P has at least one inner vertex, the inner vertices are covered if and only if there exist $(p, r) \neq (q, s) \in \mathcal{E} \times \mathcal{R}$ and a shortest p^r_S–q^s_S-path uses P (that is, $d(p^r, \ell^{\leftarrow}) + h_\ell + d(\ell^{\rightarrow}, q^s) = d(p^r, q^s)$).

(iii) For each $p \in E(\widetilde{G})$ and $r \in \mathcal{R}$, the vertices in V^r_p are covered if it holds that $x^r_p \leq 1$ (that is, $V^r_p = \emptyset$) or there exists $(q, s) \in \mathcal{E} \times \mathcal{R}$ such that a shortest p^r_S–q^s_S-path visits p^r.

For each of the cases (i)–(iii) and each of the parts to cover, the algorithm guesses $(p, r) \in \mathcal{E} \times \mathcal{R}$ (and $(q, s) \in \mathcal{E} \times \mathcal{R}$); recall that $|\mathcal{E}| \leq |E(\widetilde{G})| \leq 3\,\mathrm{fen}(G) - 3$ and $|V(\widetilde{G})| \leq 2\,\mathrm{fen}(G) - 2$ and $|\mathcal{R}| = 2$, so there are at most $\mathrm{fen}(G)^{O(\mathrm{fen}(G))}$ possibilities. For each guess, our algorithm adds the corresponding constraints according to Constraint (2) or Constraint (3) and checks feasibility with the now completed ILP formulation. We show that this approach is correct.

Theorem 7 Geodetic Set *can be solved in $O^*(\mathrm{fen}(G)^{O(\mathrm{fen}(G))})$ time.*[3]

Proof: We prove that there is a geodetic set $S \supseteq T$ satisfying Lemmata 8 and 9 if and only if one of our ILP instances is a yes-instance. The forward direction is clearly correct. The correctness of the other direction follows from Observation 6.

Note that we construct $\mathrm{fen}(G)^{O(\mathrm{fen}(G))}$ instances of ILP. Each ILP instance uses $O(\mathrm{fen}(G))$ variables, so solving it takes $O^*(\mathrm{fen}(G)^{O(\mathrm{fen}(G))})$ time [20]. This results in an algorithm whose running time is $O^*(\mathrm{fen}(G)^{O(\mathrm{fen}(G))})$. □

5 Fixed-Parameter Tractability for Clique-Width with Diameter

In this section we obtain fixed-parameter tractability results for clique-width combined with diameter, and for tree-depth. Our algorithm is based on a theorem by Courcelle et al. [10]: If a graph property π can be expressed as a formula φ in MSO_1 logic, then whether a graph G has π can be determined in $O(f(\mathrm{cw}(G) + |\varphi|) \cdot (|V(G)| + |E(G)|))$ time for some function f.

[3] the $O^*(\cdot)$ notation hides factors that are polynomial in the input size

Theorem 8 Geodetic Set *is fixed-parameter tractable with respect to the clique-width and the diameter of the input graph, combined.*

Proof: We describe how to express Geodetic Set in MSO_1 logic. We define

$$\varphi = \exists S \ (\forall v \left[\exists u, w \left(u \in S \wedge w \in S \wedge \text{Visit}(u, v, w)\right)\right]),$$

where $\text{Visit}(u, v, w)$ is true if and only if there is a shortest path u–w visiting v. For the construction of $\text{Visit}(u, v, w)$ let us first define a formula $\text{Path}(v_1, \dots, v_i)$ which evaluates to true if and only if $(v_1, \dots, v_i)$ is a path:

$$\text{Path}(v_1, \dots, v_\delta) = \bigwedge_{j \in [i-1]} v_j v_{j+1} \in E(G).$$

We then define $\text{Dist}_i(u, w)$ which is true if and only if $d_G(u, w) = i$.

$$\begin{aligned}\text{Dist}_i(u, w) = {} & \exists v_2, \dots, v_{i-1} \left(\text{Path}(u, v_2, \dots, v_{i-1}, w)\right) \\ & \wedge \bigwedge_{j \in [i-1]} \nexists v_2, \dots, v_{j-1} (\text{Path}(u, v_2, \dots, v_{j-1}, w)).\end{aligned}$$

Finally, we define $\text{Visit}(u, v, w)$:

$$\text{Visit}(u, v, w) = \bigvee_{i \in [\text{diam}(G)]} \left(\text{Dist}_i(u, w) \wedge \left[\bigvee_{j \in [i-1]} \text{Dist}_j(u, v) \wedge \text{Dist}_{j-i}(v, w) \right] \right).$$

Note that $|\varphi| \in \text{diam}(G)^{O(1)}$. Thus, fixed-parameter tractability for $\text{cw}(G) + \text{diam}(G)$ follows from Courcelle's theorem. □

Note that $\text{cw}(G) \leq 2$ and $\text{diam}(G) \leq 2$ for any cograph G. Thus, our result extends polynomial-time solvability on cographs proven by Dourado et al. [12].

We also obtain fixed-parameter tractability for tree-depth as well as for modular-width from Theorem 8. The tree-depth of a graph G can be roughly approximated by $\log h \leq \text{td}(G) \leq h$, where h is the height of a depth-first search tree of G [22]. Hence, the length of all paths in G, specifically the diameter of G, is at most $2^{\text{td}(G)}$. Moreover, $\text{cw}(G) \leq 3 \cdot 2^{\text{tw}(G)-1}$ [9] and $\text{tw}(G) \leq \text{td}(G) - 1$. Similarly, $\text{cw}(G) \leq \text{mw}(G)$ (by definition) and $\text{diam}(G) \leq \max\{2, \text{mw}(G)\}$ [19]. Consequently, we obtain the following.

Corollary 9 Geodetic Set *is fixed-parameter tractable with respect to tree-depth and with respect to modular-width.*

6 Conclusion

We initiated a parameterized complexity study of Geodetic Set for parameters measuring tree-likeness. We conclude this work by suggesting some future research directions. None of the fixed-parameter algorithms presented in this work are practical. Are there more efficient fixed-parameter algorithms with respect to feedback edge number, tree-depth or modular-width? Further, while we can quite surely exclude fixed-parameter tractability for feedback vertex number and path-width, it is still open whether Geodetic Set is in XP with any (combination) of these parameters.

Recall that the related Geodetic Hull problem is in XP with respect to tree-width [18], but for Geodetic Set, even the complexity on series-parallel graphs (which have tree-width two) is unknown.

Going to related problems and parameters, it is open whether Metric Dimension is fixed-parameter tractable with respect to the feedback edge number [14]. This is especially interesting since the problem behaves similarly to Geodetic Set in terms of complexity: Metric Dimension is fixed-parameter tractable with respect to tree-depth [23] and with respect to modular-width [3], but W[1]-hard with respect to path-width [5] and W[2]-hard with respect to the solution size [17]. We are optimistic that the method presented in Section 4 can be used to answer this question positively, especially since Epstein et al. [15] showed that the number of solution vertices on a path of degree-two vertices (cf. Lemma 9) is bounded by a constant.

Acknowledgements. We thank Lucia Draque Penso (Ulm University) for suggesting studying Geodetic Set from a view of parameterized complexity, and we thank André Nichterlein and Rolf Niedermeier (both TU Berlin) for helpful feedback and discussion. We are also grateful to anonymous reviewers for suggesting that the ILP instances in Section 4 can be solved more efficiently.

References

[1] Júlio Araújo, Grégory Morel, Leonardo Sampaio, Ronan Pardo Soares, and Valentin Weber. Hull number: P_5-free graphs and reduction rules. *Discrete Applied Mathematics*, 210:171–175, 2016. doi:10.1016/j.dam.2015.03.019.

[2] Mustafa Atici. Computational complexity of geodetic set. *International Journal of Computer Mathematics*, 79(5):587–591, 2002. doi:10.1080/00207160210954.

[3] Rémy Belmonte, Fedor V. Fomin, Petr A. Golovach, and M. S. Ramanujan. Metric dimension of bounded tree-length graphs. *SIAM Journal on Discrete Mathematics*, 31(2):1217–1243, 2017. doi:10.1137/16M1057383.

[4] Stéphane Bessy, Mitre Costa Dourado, Lucia Draque Penso, and Dieter Rautenbach. The geodetic hull number is hard for chordal graphs. *SIAM Journal on Discrete Mathematics*, 32(1):543–547, 2018. doi:10.1137/17M1131726.

[5] Édouard Bonnet and Nidhi Purohit. Metric dimension parameterized by treewidth. In *14th International Symposium on Parameterized and Exact Computation (IPEC '19)*, pages 5:1–5:15, 2019. doi:10.1007/s00453-021-00808-9.

[6] Bostjan Bresar, Sandi Klavzar, and Aleksandra Tepeh Horvat. On the geodetic number and related metric sets in Cartesian product graphs. *Discrete Mathematics*, 308(23):5555–5561, 2008. doi:10.1016/j.disc.2007.10.007.

[7] Letícia Rodrigues Bueno, Lucia Draque Penso, Fábio Protti, Victor R. Ramos, Dieter Rautenbach, and Uéverton S. Souza. On the hardness of finding the geodetic number of a subcubic graph. *Information Processing Letters*, 135:22–27, 2018. doi:10.1016/j.ipl.2018.02.012.

[8] Dibyayan Chakraborty, Florent Foucaud, Harmender Gahlawat, Subir Kumar Ghosh, and Bodhayan Roy. Hardness and approximation for the geodetic set problem in some graph

classes. In *6th International Conference on Algorithms and Discrete Applied Mathematics (CALDAM '20)*, pages 102–115, 2020. doi:10.1007/978-3-030-39219-2_9.

[9] Derek G. Corneil and Udi Rotics. On the relationship between clique-width and treewidth. *SIAM Journal on Computing*, 34(4):825–847, 2005. doi:10.1137/S0097539701385351.

[10] Bruno Courcelle, Johann A. Makowsky, and Udi Rotics. Linear time solvable optimization problems on graphs of bounded clique-width. *Theory of Computing Systems*, 33(2):125–150, 2000. doi:10.1007/s002249910009.

[11] Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015. doi:10.1007/978-3-319-21275-3.

[12] Mitre Costa Dourado, Fábio Protti, Dieter Rautenbach, and Jayme Luiz Szwarcfiter. Some remarks on the geodetic number of a graph. *Discrete Mathematics*, 310(4):832–837, 2010. doi:10.1016/j.disc.2009.09.018.

[13] Mitre Costa Dourado, Lucia Draque Penso Rautenbach, and Dieter Rautenbach. On the geodetic hull number of P_k-free graphs. *Theoretical Computer Science*, 640:52–60, 2016. doi:10.1016/j.tcs.2016.05.047.

[14] David Eppstein. Metric dimension parameterized by max leaf number. *Journal of Graph Algorithms and Applications*, 19(1):313–323, 2015. doi:10.7155/jgaa.00360.

[15] Leah Epstein, Asaf Levin, and Gerhard J. Woeginger. The (weighted) metric dimension of graphs: Hard and easy cases. *Algorithmica*, 72(4):1130–1171, 2015. doi:10.1007/s00453-014-9896-2.

[16] Michel Habib and Christophe Paul. A survey of the algorithmic aspects of modular decomposition. *Computer Science Review*, 4(1):41–59, 2010. doi:10.1016/j.cosrev.2010.01.001.

[17] Sepp Hartung and André Nichterlein. On the parameterized and approximation hardness of metric dimension. In *Proceedings of the 28th IEEE Conference on Computational Complexity (CCC '13)*, pages 266–276. IEEE, 2013. doi:10.1109/CCC.2013.36.

[18] Mamadou Moustapha Kanté, Thiago Braga Marcilon, and Rudini M. Sampaio. On the parameterized complexity of the geodesic hull number. *Theoretical Computer Science*, 791:10–27, 2019. doi:10.1016/j.tcs.2019.05.005.

[19] Minki Kim, Bernard Lidický, Tomás Masarík, and Florian Pfender. Notes on complexity of packing coloring. *Information Processing Letters*, 137:6–10, 2018. doi:10.1016/j.ipl.2018.04.012.

[20] Hendrik W. Lenstra. Integer programming with a fixed number of variables. *Mathematics of Operations Research*, 8:538–548, 1983. doi:10.1287/moor.8.4.538.

[21] Dániel Marx. On the optimality of planar and geometric approximation schemes. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS '07)*, pages 338–348, 2007. doi:10.1109/FOCS.2007.50.

[22] Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity: Graphs, Structures, and Algorithms*, volume 28 of *Algorithms and Combinatorics*. Springer, 2012. doi:10.1007/978-3-642-27875-4.

[23] Sanchez Villaamil. *About Treedepth and Related Notions*. PhD thesis, RWTH Aachen, 2017. doi:10.18154/RWTH-2017-09829.