Abstract

Treewidth serves as an important parameter that, when bounded, yields tractability for a wide class of problems. For example, graph problems expressible in Monadic Second Order (MSO) logic and Quantified SAT or, more generally, Quantified CSP, are fixed-parameter tractable parameterized by the treewidth of the input’s (primal) graph plus the length of the MSO-formula [Courcelle, Information & Computation 1990] and the quantifier rank [Chen, ECAI 2004], respectively. The algorithms generated by these (meta-)results have running times whose dependence on treewidth is a tower of exponents. A conditional lower bound by Fichte, Hecher, and Pfandler [LICS 2020] shows that, for Quantified SAT, the height of this tower is equal to the number of quantifier alternations. These types of lower bounds, which show that at least double-exponential factors in the running time are necessary, exhibit the extraordinary level of computational hardness for such problems, and are rare in the current literature: there are only a handful of such lower bounds (for treewidth and vertex cover parameterizations) and all of them are for problems that are $\mathbf{NP}$-complete, $\mathbf{\Sigma}_p^2$-complete, $\mathbf{\Pi}_p^2$-complete, or complete for even higher levels of the polynomial hierarchy.

Our results demonstrate, for the first time, that it is not necessary to go higher up in the polynomial hierarchy to achieve double-exponential lower bounds: we derive double-exponential lower bounds in the treewidth ($tw$) and the vertex cover number ($vc$), for natural, important, and well-studied $\mathbf{NP}$-complete graph problems. Specifically, we design a technique to obtain such lower bounds and show its versatility by applying it to three different problems: Metric Dimension, Strong Metric Dimension, and Geodetic Set. We prove that these problems do not admit $2^{2^{O(tw)}} \cdot n^{O(1)}$-time algorithms, even on bounded diameter graphs, unless the ETH fails (here, $n$ is the number of vertices in the graph). In fact, for Strong Metric Dimension, the double-exponential lower bound holds even for the vertex cover number. We further complement all our lower bounds with matching (and sometimes non-trivial) upper bounds.

For the conditional lower bounds, we design and use a novel, yet simple technique based on Sperner families of sets. We believe that the amenability of our technique will lead to obtaining such lower bounds for many other problems in $\mathbf{NP}$. 
Introduction

Many interesting computational problems turn out to be intractable. In these cases, identifying parameters under which the problems become tractable is desirable. In the area of parameterized complexity, treewidth is a cornerstone parameter since a large class of problems become tractable on graphs of bounded treewidth.

Courcelle’s celebrated theorem [13] states that the class of graph problems expressible in Monadic Second-Order Logic (MSOL) of constant size is fixed-parameter tractable (FPT) when parameterized by the treewidth of the graph. That is, such problems admit algorithms whose running time is of the form $f(tw) \cdot \text{poly}(n)$, where $tw$ is the treewidth of the input, $n$ is the size of the input, and $f$ is a function that depends only on $tw$. Similarly, a result by Chen [12] shows that the QUANTIFIED SAT (Q-SAT) problem can also be solved in time $f(tw) \cdot \text{poly}(n)$, where $tw$ is the treewidth of the primal graph of the input formula and $f$ is a function that depends only on $tw$ and the number of quantifier alternations in the input formula. Q-SAT is a generalization of SAT that allows universal and existential quantifications over the variables. Note that Q-SAT with $k$ quantifier alternations is $\Pi_k^P$-complete or $\Sigma_k^P$-complete. Unfortunately, in both of the aforementioned results, the function $f$ is a tower of exponents whose height depends roughly on the size of the MSOL and input formulas, respectively. For Q-SAT, the height of this tower equals the number of quantifier alternations in the Q-SAT instance [12].

Over the years, the focus shifted to making such FPT algorithms as efficient as possible. Thus, a natural question is to ask when this higher-exponential dependence on treewidth is necessary. There is a rich literature that provides (conditional) lower bounds on this dependency for many problems, and these bounds are commonly of the form $2^{o(tw)}$ or, in some unusual cases, $2^{o(tw \log tw)}$ (e.g., [15, 50]) and even $2^{o(\text{poly}(tw))}$ (e.g., [14, 55]). Most notably, these lower bounds are far from the tower of exponents upper bounds given by the (meta-)results discussed above. In this work, we develop a simple technique that allows to prove double-exponential dependence on the treewidth $tw$ and the vertex cover number $vc$.
two of the most fundamental graph parameters. Notably, these are the first such results for problems in \( \mathrm{NP} \), and we believe that the amenability of our technique will lead to many more similar results for other problems in \( \mathrm{NP} \).

Indeed, after a preprint of this paper appeared on arxiv, our technique was also used to prove double-exponential dependence on \( \nu \mathrm{c} \) for an \( \mathrm{NP} \)-complete machine learning problem \([11]\) and double-exponential dependence on the solution size and \( \nu \mathrm{w} \) for \( \mathrm{NP} \)-complete identification problems like Test Cover and Locating-Dominating Set \([10]\).

**Double-exponential lower bounds: treewidth and vertex cover parameterizations.** Fichte, Hecher, and Pfander \([25]\) recently proved that, assuming the Exponential Time Hypothesis\(^1\) (\(\text{ETH}\)), Q-SAT with \(k\) quantifier alternations cannot be solved in time significantly better than a tower of exponents of height \(k\) in the treewidth. This exemplifies an interesting but expected trait of this problem: its complexity, in terms of the height of the exponential tower in \( \nu \mathrm{w} \), increases with each quantifier alternation. It strengthened the result that appeared in \([48]\), where conditional double-exponential lower bounds for \( \exists \forall \mathrm{SAT} \) and \( \forall \exists \mathrm{SAT} \) were given. The results in \([48]\) also yield a double-exponential lower bound in \( \nu \mathrm{c} \) of the primal graph for both problems. Besides these results, there are only a handful of other problems known to require higher-exponential dependence in the treewidth of the input graph (or the primal graph of the input formula). Specifically, the \( \Pi^p_2 \)-complete \( k \)-CHOOSABILITY problem and the \( \Sigma^p_2 \)-complete \( k \)-CHOOSABILITY DELETION problem admit a double-exponential and a triple-exponential lower bound in treewidth \([52]\), respectively. Recently, the \( \Sigma^p_2 \)-complete problems CYCLE HitPACK and \( H \)-HitPACK, for a fixed graph \( H \), were shown to admit tight algorithms that are double-exponential in the treewidth \([26]\). Further, the \( \Sigma^p_2 \)-complete problem CORE STABILITY was shown to admit a tight double-exponential lower bound in the treewidth, even on graphs of bounded degree \([32]\). Lastly, the \#\( \mathrm{NP} \)-complete counting problem PROJECTED MODEL COUNTING admits a double-exponential lower bound in \( \nu \mathrm{w} \) \([23, 24]\). For other double-exponential lower bounds, see \([1, 16, 27, 32, 37, 41, 44, 47, 51, 56, 59]\).

All the double- (or higher) exponential lower bounds in treewidth mentioned so far are for problems that are \#\( \mathrm{NP} \)-complete, \( \Sigma^p_2 \)-complete, \( \Pi^p_2 \)-complete, or complete for even higher levels of the polynomial hierarchy. To quote \([52]\): “\( \Pi^p_2 \)-completeness of these problems already gives sufficient explanation why double- [...] exponential dependence on treewidth is needed. [...] the quantifier alternations in the problem definitions are the common underlying reasons for being in the higher levels of the polynomial hierarchy and for requiring unusually large dependence on treewidth.”

As mentioned above, we develop a technique that allows to demonstrate, for the first time, that it is not necessary to go to higher levels of the polynomial hierarchy to achieve double-exponential lower bounds in the treewidth or the vertex cover number of the graph.

---

1. The Exponential Time Hypothesis roughly states that \( n \)-variable 3-SAT cannot be solved in time \( 2^{o(n)} \).
2. While it may be possible to artificially engineer a graph problem or graph representation of a problem in \( \mathrm{NP} \) that admits such lower bounds (although, to the best of our knowledge, this has not been done), we emphasize that this is not the case for these three natural and well-established graph problems in \( \mathrm{NP} \).
NP-complete metric-based graph problems. We study three metric-based graph problems. These problems are Metric Dimension [34, 58], Strong Metric Dimension [57], and Geodetic Set [33], and they arise from network design and network monitoring. Apart from serving as examples for double-exponential dependence on treewidth and the amenability of our technique, these problems are of interest in their own right, and possess a rich literature both in the algorithms and discrete mathematics communities (see, e.g., [2, 3, 4, 5, 7, 8, 17, 19, 20, 21, 31, 38, 45, 53] and the references below). Their non-local nature has posed interesting algorithmic challenges and our results, as we explain later, supplement the already vast literature on the structural parameterizations of these problems. Below we define the three above-mentioned problems formally, and particularly focus on Metric Dimension as it is the most popular and well-studied of the three.

**Metric Dimension**

**Input:** A graph $G$ and a positive integer $k$.

**Question:** Does there exist $S \subseteq V(G)$ such that $|S| \leq k$ and, for any pair of vertices $u, v \in V(G)$, there exists a vertex $w \in S$ with $d(w, u) \neq d(w, v)$?

The Metric Dimension problem dates back to the 70s [34, 58]. As in geolocation problems, the aim is to distinguish the vertices of a graph via their distances to a solution set. Metric Dimension was first shown to be NP-complete in general graphs in Garey and Johnson’s book [30, GT61], and this was later extended to many restricted graph classes [18, 22, 28], including graphs of diameter 2 [28] and graphs of pathwidth 24 [49]. In a seminal paper, Metric Dimension was proven to be W[2]-hard parameterized by the solution size $k$, even in subcubic bipartite graphs [35]. This drove the subsequent meticulous study of the problem under structural parameterizations.

In particular, the complexity of Metric Dimension parameterized by treewidth remained an intriguing open problem for a long time. Recently, it was shown that Metric Dimension is para-NP-hard parameterized by pathwidth (pw) [49] (an earlier result [6] showed that it is W[1]-hard for pathwidth). A subsequent paper showed that the problem is W[1]-hard parameterized by the combined parameter feedback vertex set number (fvs) plus pathwidth of the graph [29].

We conclude this part with the definitions of the remaining two problems, both of which are known to be NP-Complete [9, 54]. Geodetic Set is also W[1]-hard parameterized by the solution size, feedback vertex set number, and pathwidth, combined [39].

**Strong Metric Dimension**

**Input:** A graph $G$ and a positive integer $k$.

**Question:** Does there exist $S \subseteq V(G)$ such that $|S| \leq k$ and, for any pair of vertices $u, v \in V(G)$, there exists a vertex $w \in S$ such that either $u$ lies on some shortest path between $v$ and $w$, or $v$ lies on some shortest path between $u$ and $w$?

**Geodetic Set**

**Input:** A graph $G$ and a positive integer $k$.

**Question:** Does there exist $S \subseteq V(G)$ such that $|S| \leq k$ and, for any vertex $u \in V(G)$, there are two vertices $s_1, s_2 \in S$ such that a shortest path from $s_1$ to $s_2$ contains $u$?

Our technical contributions. As Metric Dimension and Geodetic Set are NP-complete on bounded diameter graphs or on bounded treewidth graphs, we study their parameterized complexity with $tw + diam$ as the parameter and prove the following results.
1. **Metric Dimension** and **Geodetic Set** do not admit algorithms running in time $2^{f(\text{diam})^{O(1)}} \cdot n^{O(1)}$, for any computable function $f$, unless the ETH fails.

2. **Strong Metric Dimension** does not even admit an algorithm with a running time of $2^{2^{O(\text{vc})}} \cdot n^{O(1)}$, unless the ETH fails. This also implies the problem does not admit a kernelization algorithm that outputs an instance with $2^{2^{O(\text{vc})}}$ vertices, unless the ETH fails.

The above lower bounds for $\text{tw} + \text{diam}$, in particular, imply that **Metric Dimension** and **Geodetic Set** on graphs of bounded diameter cannot admit $2^{2^{O(\text{vc})}} \cdot n^{O(1)}$-time algorithms, unless the ETH fails. The reduction for **Metric Dimension** (sketched in Section 2.3) also works for **FVS** and **td**, and the one for **Geodetic Set** (Section 3) also works for **td**.

We show that all our lower bounds are tight by providing algorithms (kernelization algorithms, respectively) with matching running times (guarantees, respectively).

1. **Metric Dimension** and **Geodetic Set** admit algorithms running in time $2^{\text{diam}^{O(\text{tw})}} \cdot n^{O(1)}$.

2. **Strong Metric Dimension** admits an algorithm running in time $2^{2^{O(\text{vc})}} \cdot n^{O(1)}$ and a kernel with $2^{O(\text{vc})}$ vertices.

The (kernelization) algorithm for the $\text{vc}$ parameterization is very simple, whereas the algorithms for the $\text{tw} + \text{diam}$ parameter are highly non-trivial and require showing interesting locality properties in the instance. Further, for our $\text{tw} + \text{diam}$ parameterized algorithms, the (double-exponential) dependency of treewidth in the running time is unusual (and rightly so, as exhibited by our lower bounds), as most natural graph problems in NP for which a dedicated algorithm (i.e., not relying on Courcelle’s theorem) parameterized by treewidth is known, can be solved in time $2^{O(\text{tw})} \cdot n^{O(1)}$, $2^{O(\text{tw} \log(\text{tw}))} \cdot n^{O(1)}$ or $2^{O(\text{poly}(\text{tw}))} \cdot n^{O(1)}$.

Finally, our reductions rely on a novel, yet simple technique based on Sperner families of sets that allows to encode particular SAT relations across large sets of variables and clauses into relatively small vertex-separators. As mentioned before, we believe that this technique is the key to obtaining such lower bound results for other problems in NP. In particular, as witnessed by our results, our technique has the additional features that it even allows to prove such lower bounds in very restricted cases, such as bounded diameter graphs, and is not specific to any one structural parameter, as it also works for, e.g., the feedback vertex set number and treedepth.

Due to space constraints, we cannot discuss all of our results in depth, and refer the reader to the full version for full proofs, formal details, and more related work. Nonetheless, we elaborate on our technique and present an overview of the results for **Metric Dimension** in the next section. Then, in Section 3, we present a formal proof for the lower bound for **Geodetic Set**. Finally, we conclude the paper in Section 4.

## 2 Technical Overview

In this section, we present an overview of our lower bound techniques. We first exhibit our technique to obtain the double-exponential lower bounds in its most general setting. Then, we continue with the problem-specific tools we developed that are required for the reductions.
Figure 1 Graph representations of $3$-Partitioned-$3$-SAT. (Left) incidence graph representation. (Right) representation with small separators using our technique. Note, for example, that $x^1_1$ appears as a positive literal in the clause $C_1$. Thus, on the left, $t^1_2$ is the only literal vertex in $A^\alpha$ incident to $c_1$, while on the right, $t^\beta_2$ is the only literal vertex in $A^\alpha$ that does not share a common neighbor with $c_1$ in $V^\alpha$. The edges from $c_2$ to each vertex in $V^\alpha$ are omitted for clarity.

### 2.1 General Technique for Double-Exponential Lower Bounds

The first integral part of our technique is to reduce from a variant of $3$-SAT known as $3$-Partitioned-$3$-SAT that was introduced in [46]. In this problem, the input is a formula $\psi$ in $3$-CNF form, together with a partition of the set of its variables into three disjoint sets $X^\alpha$, $X^\beta$, $X^\gamma$, with $|X^\alpha| = |X^\beta| = |X^\gamma| = n$, and such that no clause contains more than one variable from each of $X^\alpha$, $X^\beta$, and $X^\gamma$. The objective is to determine whether $\psi$ is satisfiable. Unless the ETH fails, $3$-Partitioned-$3$-SAT does not admit an algorithm running in time $2^{o(n)}$ [46, Theorem 3].

Typical reductions from satisfiability problems to graph problems usually entail representing the satisfiability problem by its incidence graph, in which each variable is represented by two vertices corresponding to its positive and negative literals. In this representation, a clause vertex is adjacent to a literal vertex if and only if it contains that literal in $\psi$ (see Figure 1 (left) for an illustration). However, this naive approach does not lead to any structural parameters of the incidence graph being of bounded size. The core idea of our technique is to instead represent the relationships between clause and literal vertices via edges from these two sets of vertices to “small” separators (three separators in the case of $3$-Partitioned-$3$-SAT) that encode these relationships.

Formally, this is achieved as follows. For a positive integer $p$, define $F_p$ as the collection of subsets of $[2p]$ that contains exactly $p$ integers. We critically use the fact that no set in $F_p$ is contained in any other set in $F_p$ (such a collection of sets are called a Sperner family). Let $\ell$ be a positive integer such that $\ell \leq \binom{2p}{p}$. We define $\text{set-rep} : [\ell] \mapsto F_p$ as a one-to-one function by arbitrarily assigning a set in $F_p$ to an integer in $[\ell]$. By the asymptotic estimation of the central binomial coefficient, $\binom{2p}{p} \sim \frac{4^p}{\sqrt{\pi p}}$ [36]. To get the upper bound of $p$, we scale down the asymptotic function and have $\ell \leq \frac{4^p}{\sqrt{\pi p}} = 2^p$. Thus, $p = O(\log \ell)$.

Let $\psi$ be an instance of $3$-Partitioned-$3$-SAT on $3n$ variables, and let $p$ be the smallest integer such that $2n \leq \binom{2p}{p}$. In particular, $p = O(\log n)$. Define $\text{set-rep} : [2n] \mapsto F_p$ as above. Rename the variables in $X^\alpha$ to $x^\alpha_i$ for all $i \in [n]$. For each variable $x^\alpha_i$, add two vertices $t^1_2$.
and $f_{2i-1}^α$ corresponding to the positive and negative literals of $x_i^α$, respectively. Let $A^α = \{t^α_{2i}, f^α_{2i-1} \mid i \in [n]\}$. Add a validation portal with $2p$ vertices, denoted by $V^α = \{v^α_1, \ldots, v^α_{2p}\}$. For each $i \in [n]$, add the edge $t^α_{2i}v^α_i$ for each $p' \in set-rep(2i)$. Similarly, for each $i \in [n]$, add the edge $f^α_{2i-1}v^α_i$ for each $p' \in set-rep(2i - 1)$. Repeat the above steps for $β$ and $γ$.

Now, for each clause $C_j$ ($j \in [m]$) in $ψ$, add a clause vertex $c_j$. Let $δ \in \{α, β, γ\}$. For all $i \in [n]$ and $j \in [m]$, if the variable $x^δ_i$ appears as a positive (negative, respectively) literal in the clause $C_j$ in $ψ$, then add the edge $c_jv^δ_i$ for each $p' \in [2p] \setminus set-rep(2i)$ ($p' \in [2p] \setminus set-rep(2i - 1)$, respectively). For all $j \in [m]$, if no variable from $X^δ$ appears in $C_j$ in $ψ$, then make $c_j$ adjacent to all the vertices in $V^δ$. See Figure 1 (right) for an illustration.

As a clause contains at most one variable from $X^δ$ in $ψ$, $c_j$ and $t^δ_{2i}$ ($f^δ_{2i-1}$, respectively) do not share a common neighbor in $V^δ$ if and only if the clause $C_j$ contains $x^δ_i$ as a positive (negative, respectively) literal in $ψ$. For the reductions, we use this representation of the relationship between clause and literal vertices. Since $p = O(\log n)$, this ensures that $ω(G) = O(\log n)$, which we exploit along with the fact that, unless the ETH fails, 3-PARTITIONED-3-SAT does not admit an algorithm running in time $2^{o(n)}$.

### 2.2 Basic Tools for Lower Bounds

For brevity, we focus on METRIC DIMENSION and explain our problem-specific tools in this context. We use two such simple tools: the bit representation gadget and the set representation gadget. The set representation gadget is the problem-specific implementation of the above technique, and it uses the bit representation gadget.

Before going further, we need to define some terms related to METRIC DIMENSION. The set $S$ defined in the problem statement of METRIC DIMENSION is called a resolving set of $G$. A subset of vertices $S' \subseteq V(G)$ resolves a pair of vertices $u, v \in V(G)$ if there exists a vertex $w \in S'$ such that $d(w, u) ≠ d(w, v)$. Lastly, a vertex $u \in V(G)$ is distinguished by a subset of vertices $S' \subseteq V(G)$ if, for any $v \in V(G) \setminus \{u\}$, there exists a vertex $w \in S'$ such that $d(w, u) ≠ d(w, v)$.

**Bit Representation Gadget to Identify Sets.** Suppose we are given a graph $G'$ and a subset $X \subseteq V(G')$ of its vertices. Further, suppose that we want to add a vertex set $X^+$ to $G'$ to obtain a new graph $G$ with the following properties. We want that each vertex in $X \cup X^+$ is
distinguished by vertices in $X^+$ that must be in any resolving set $S$ of $G$, and that no vertex in $X^+$ can resolve any “critical pair” of vertices in $G$. Roughly, a pair of vertices is critical if it forces certain “types” of vertices to be in any resolving set $S$ of $G$, and the selection of the specific vertices of those types depends on the solution to the problem being reduced from (which, in our case, is 3-PARTITIONED-3-SAT [46]). We refer to the graph induced by the vertices of $X^+$, along with the edges connecting $X^+$ to $G'$, as the Set Identifying Gadget for the set $X$. Given a graph $G'$ and a non-empty subset $X \subseteq V(G')$ of its vertices, to construct such a graph $G$, we add vertices and edges to $G'$ as follows (see Figure 2):

- The vertex set $X^+$ that we are aiming to add is the union of a set bit-rep$(X)$ and a special vertex denoted by nullifier$(X)$.
- First, let $X = \{x_i \mid i \in [\lceil |X| \rceil]\}$, and set $q := \lceil \log(\lceil |X| \rceil + 2) \rceil + 1$. We select this value for $q$ to (1) uniquely represent each integer in $[|X|]$ by its bit-representation in binary (note that we start from 1 and not 0), (2) ensure that the only vertex whose bit-representation contains all 1’s is nullifier$(X)$, and (3) reserve one spot for an additional vertex $y_i$.

- For every $i \in [q]$, add three vertices $y^i_1, y^i_2, y^i_3$ and add the path $(y^i_1, y^i_2, y^i_3)$.
- Add 3 vertices $y^1, y^2, y^3$ and the path $(y^1, y^2, y^3)$. Add edges to make $\{y_i \mid i \in [q]\} \cup \{y\}$ a clique. Make $y$ adjacent to each vertex in $X$. Let bit-rep$(X) = \{y_i, y^i_1, y^i_2 \mid i \in [q]\} \cup \{y, y^1, y^2, y^3\}$ and denote its subset by $\text{bits}(X) = \{y^i_1, y^i_2 \mid i \in [q]\} \cup \{y^1, y^2, y^3\}$.
- For every integer $j \in [|X|]$, let $\text{bin}(j)$ denote the binary representation of $j$ using $q$ bits. Connect $x_j$ with $y_i$ if the $i^\text{th}$ bit (going from left to right) in $\text{bin}(j)$ is 1.
- Add a vertex, denoted by nullifier$(X)$, and connect it to each vertex in $\{y_i \mid i \in [q]\} \cup \{y\}$.
- For every vertex $u \in V(G) \setminus (X \cup X^+)$ such that $u$ is adjacent to some vertex in $X$, add an edge between $u$ and nullifier$(X)$. We add this vertex to ensure that vertices in bit-rep$(X)$ do not resolve critical pairs in $V(G)$.

Set Representation Gadget. We define set-rep : $[\ell] \mapsto \mathcal{F}_p$, as in Section 2.1, and recall that $p = O(\log \ell)$. Suppose we have a “large” collection of vertices, say $A = \{a_1, a_2, \ldots, a_r\}$, and a “large” collection of critical pairs $C = \{c^1, c^2, \ldots, c^m\}$. Moreover, we are given an injective function $\phi : [m] \mapsto [\ell]$. The objective is to design a gadget such that only $a_\phi(q) \in A$ can resolve a critical pair $\langle c^q, c^p \rangle \in C$ for any $q \in [m]$, while keeping the treewidth of this part of the graph of order $O(\log(|A|))$. With this in mind, we do the following.

- Add vertices and edges to identify the set $A$ and to add critical pairs in $C$ (for each critical pair in $C$, both vertices share the same bit-representation in the Set Identifying Gadget for $C$).

- Add a validation portal, a clique on $2p$ vertices, denoted by $V = \{v_1, v_2, \ldots, v_{2p}\}$, and vertices and edges to identify it.

- For every $i \in [\ell]$ and for every $p' \in \text{set-rep}(i)$, add the edge $(a_i, v_{p'})$.

- For every critical pair $\langle c^q, c^p \rangle$, make $c^p$ adjacent to every vertex in $V$, and add every edge of the form $(c^p, v_{p'})$ for $p' \in [2p] \setminus \text{set-rep}(\phi(q))$. Note that the vertices in $V$ that are indexed using integers in $\text{set-rep}(\phi(q))$ are not adjacent with $c^p$.

See Figure 3 for an illustration. Now, consider a critical pair $\langle c^q, c^p \rangle$ and suppose $i = \phi(q)$.

- By the construction, $N(a_i) \cap N(c^q) \neq \emptyset$, whereas $N(a_i) \cap N(c^p) = \emptyset$. Hence, $a_i$ resolves the critical pair $\langle c^q, c^p \rangle$ as $d(a_i, c^q) = 2$ and $d(a_i, c^p) > 2$.

- For any other vertex in $A$, say $a_j$, set-rep$(j) \setminus \text{set-rep}(i)$ is a non-empty set. So, there are paths from $a_j$ to $c^q$ and $a_j$ to $c^p$ through vertices in $V$ with indices in set-rep$(j) \setminus \text{set-rep}(i)$. This implies that $d(a_j, c^q) = d(a_j, c^p) = 2$ and $a_j$ cannot resolve the pair $\langle c^q, c^p \rangle$. 


2.3 Sketch of the Lower Bound Proof for Metric Dimension

With these tools in hand, we present an overview of the reduction from 3-PARTITIONED-3-SAT used to prove Theorem 1, which we restate here for convenience.

\textbf{Theorem 1.} Unless the ETH fails, Metric Dimension does not admit an algorithm running in time \(2^{f(\text{diam}(G))} \cdot n^{O(1)}\) for any computable function \(f: \mathbb{N} \rightarrow \mathbb{N}\).

The reduction in the proof of Theorem 1 takes as input an instance \(\psi\) of 3-PARTITIONED-3-SAT on \(3n\) variables and returns \((G, k)\) as an instance of Metric Dimension such that \(\text{tw}(G) = O(\log(n))\) and \(\text{diam}(G) = O(1)\). In the following, we mention a crude outline of the reduction, omitting some technical details.

2.3.1 Reduction

- We rename the variables in \(X^o\) to \(x_i^o\) for \(i \in [n]\). For every variable \(x_i^o\), we add a critical pair \(\langle x_i^o, x_i^{o,*} \rangle\) of vertices. We denote \(X^o = \{x_i^o, x_i^{o,*} \mid i \in [n]\}\).

- For each variable \(x_i^o\), we add the vertices \(t_{2i}^o, f_{2i-1}^o\) and \(f_{2i}^o, t_{2i-1}^o\). Let \(A^o = \{t_{2i}^o, f_{2i-1}^o \mid i \in [n]\}\).

- For every \(i \in [n]\), we add the edges \((x_i^o, t_{2i}^o)\) and \((x_i^o, f_{2i-1}^o)\) which will ensure that any resolving set contains at least one vertex in \(\{t_{2i}^o, f_{2i-1}^o, x_i^o, x_i^{o,*}\}\) for every \(i \in [n]\).

- Let \(p\) be the smallest integer such that \(2n \leq \binom{2p}{p}\). In particular, \(p = O(\log n)\). Define \(\text{set-rep} : [2n] \rightarrow F_p\) as in Section 2.1.

- We add a validation portal, a clique on \(2p\) vertices, denoted by \(V^o = \{v_p^o, v_2^o, \ldots, v_{2p}^o\}\).

- For each \(i \in [n]\), we add the edge \((t_{2i}^o, v_p^o)\) for every \(p' \in \text{set-rep}(2i)\). Similarly, for each \(i \in [n]\), we add the edge \((f_{2i-1}^o, v_{2p}^o)\) for every \(p' \in \text{set-rep}(2i-1)\).

We repeat the above steps to construct \(X^\beta, A^\beta, V^\beta, X^\gamma, A^\gamma, V^\gamma\).

- For every clause \(C_q\) in \(\psi\), we introduce a pair \(\langle c_q^o, c_q^{o,*} \rangle\) of vertices. Let \(C\) be the collection of vertices in such pairs.

- We add edges across \(C\) and the portals as follows. Consider a clause \(C_q\) in \(\psi\) and the corresponding critical pair \(\langle c_q^o, c_q^{o,*} \rangle\) in \(C\). Let \(\delta \in \{\alpha, \beta, \gamma\}\). As \(\psi\) is an instance of 3-PARTITIONED-3-SAT, at most one variable in \(X^\delta\) appears in \(C_q\), say \(x_i^\delta\) for some \(i \in [n]\).

We add all edges of the form \(\langle v_p^\delta, c_q^o \rangle\) for every \(p' \in [2p]\). If \(x_i^\delta\) appears as a positive literal
Figure 4 Reduction for proof of Theorem 1. Yellow lines represent that vertex is connected to every vertex in the set the edge goes to. Green edges denote adjacencies with respect to set-rep, e.g., \( t'_{2i} \) is adjacent to \( v_i \in V^\alpha \) if \( j \in \text{set-rep}(2i) \). Purple lines also indicate adjacencies with respect to set-rep, but in a complementary way, i.e., if \( x_i \in c_q \), then, for every \( p' \in [2p] \setminus \text{set-rep}(2i) \), we have \( (v'_{p'}, c'_q) \in E(G) \), and if \( \tau \in c_q \), then, for all \( p' \in [2p] \setminus \text{set-rep}(2i - 1) \), we have \( (v'_{p'}, c'_q) \in E(G) \).

In \( C_q \), then we add the edge \( (v'_{p'}, c'_q) \) for every \( p' \in [2p] \setminus \text{set-rep}(2i) \) (which corresponds to \( t^1_{2i} \)). If \( x_i^\delta \) appears as a negative literal in \( C_q \), then we add the edge \( (v'_{p'}, c'_q) \) for every \( p' \in [2p] \setminus \text{set-rep}(2i - 1) \) (which corresponds to \( f^k_{2i - 1} \)). Note that if \( x_i^\delta \) appears as a positive (negative, respectively) literal in \( C_q \), then the vertices in \( V^\delta \) whose indices are in \( \text{set-rep}(2i) \) (\( \text{set-rep}(2i - 1) \), respectively) are not adjacent to \( c'_q \). If no variable in \( X^\delta \) appears in \( C_q \), then we make each vertex in \( V^\delta \) adjacent to both \( c'_q \) and \( c^*_{q'} \).

For all the sets mentioned above, we add vertices and edges to identify them as shown in Figure 4 (for each critical pair, both vertices share the same bit-representation in their Set Identifying Gadget). This concludes the construction of \( G \). The reduction returns \( (G, k) \) as an instance of METRIC DIMENSION for some appropriate value of \( k \).

2.3.2 Correctness of the Reduction

We give an informal description of the proof of correctness of the reverse direction here. Fix \( \delta \in \{ \alpha, \beta, \gamma \} \). For all \( i \in [n] \), the only vertices that can resolve the critical pair \( (x^\delta_i, x^{\delta*}_i) \) are the vertices in \( \{ x^\delta_i, x^{\delta*}_i \} \cup \{ t^k_{2i}, f^k_{2i - 1} \} \). This fact and the budget \( k \) ensure that any resolving set of \( G \) contains exactly one vertex from \( \{ t^k_{2i}, f^k_{2i - 1} \} \cup \{ x^\delta_i, x^{\delta*}_i \} \) for all \( i \in [n] \). This naturally corresponds to an assignment of the variable \( x^\delta_i \) if a vertex from \( \{ t^k_{2i}, f^k_{2i - 1} \} \) is in the resolving set. However, if a vertex from \( \{ x^\delta_i, x^{\delta*}_i \} \) is in the resolving set, then we can see this as giving an arbitrary assignment to the variable \( x^\delta_i \). Suppose the clause \( C_q \) contains the variable \( x^\delta_i \) as a positive literal. By the construction, every vertex in \( V^\delta \) that is adjacent to \( t^k_{2i} \) is not adjacent to \( c^*_{q'} \). However, \( c^*_q \) is adjacent to every vertex in \( V^\delta \). Hence, \( d(t^k_{2i}, c^*_q) = 2 \), whereas \( d(t^k_{2i}, c^*_{q'}) > 2 \). Thus, \( t^k_{2i} \) resolves the critical pair \( (c^*_q, c^*_{q'}) \). Consider any other vertex in \( A^\delta \), say \( t^k_{2j} \). Since set-rep(2i) is not a subset of set-rep(2j) (as both have the same cardinality), there is at least one integer, say \( p' \), in set-rep(2j) \setminus \text{set-rep}(2i). The vertex \( v^p'_{p'} \in V^\delta \) is adjacent to \( t^k_{2j}, c^*_q \), and \( c^*_{q'} \). Hence, \( t^k_{2j} \) cannot resolve the critical pair \( (c^*_q, c^*_{q'}) \) as both these vertices are at distance 2 from it. Also, as \( \psi \) is an instance of
Thus, for this purpose, it is irrelevant whether which makes the proof a little bit simpler.

This also helps to encode the fact that at most one vertex from \(A^3\) should be able to resolve the critical pair \((c^3_q, c^3_q')\). Since vertices in \(X^3\) cannot resolve critical pairs \((c^3_q, c^3_q')\) in \(C\), then finding a resolving set in \(G\) corresponds to finding a satisfying assignment for \(\psi\).

### 2.3.3 Lower Bounds Obtained from the Reduction

Let \(Z = \{V^\delta \cup X^+ | X \in \{X^\delta, A^\delta, V^\delta, C\}, \delta \in \{\alpha, \beta, \gamma\}\}\). Note that \(|Z| = O(\log(n))\) and \(G - Z\) is a collection of \(P_3\)'s and isolated vertices. Hence, \(tw(G), fvs(G),\) and \(td(G)\) are upper bounded by \(O(\log(n))\). Also, \(G\) has constant diameter. Thus, if there is an algorithm for METRIC DIMENSION that runs in time \(2^{f(\text{diam}(G))} (\text{or } 2^{f(\text{diam}(G))} + 1)\), then there is an algorithm solving 3-PARTITIONED-3-SAT in time \(2^{O(n)}\), contradicting the ETH.

### 2.4 High-Level Description of the Dynamic Programming Algorithm for Metric Dimension

The aim of this subsection is to give an informal description of how we prove the upper bound concerning the parameter \(tw + \text{diam}\) for METRIC DIMENSION. To this end, we give a dynamic programming algorithm on a tree decomposition for METRIC DIMENSION. The algorithm is inspired by the one from [7] for chordal graphs, though there are some non-trivial differences. We will assume that a tree decomposition of the input graph \(G\) of width \(w\) is given to us. Note that one can compute a tree decomposition of width \(w \leq tw(G) + 1\) in time \(2^{O(tw(G))} n^{42}\), and it can be transformed into a nice tree decomposition of the same width with \(O(wn)\) bags in time \(O(w^2n)\) [40]. We now give a high-level overview of the dynamic programming algorithm used to prove the following theorem.

---

**Theorem 2.** METRIC DIMENSION admits an algorithm running in time \(2^{\text{diam}O(w)} \cdot n^{O(1)}\).

---

In [7], as the diameter of the graph was unbounded, it was crucial to restrict the computations for each step of the dynamic programming to vertices “not too far” from the current bag. This was possible due to the metric properties of chordal graphs. In our case, as we consider the diameter of the graph as a parameter, we do not need such restrictions, which makes the proof a little bit simpler.

We now give an intuitive description of the dynamic programming scheme. At each step of the algorithm, we consider a bounded number of solution types, depending on the properties of the solution vertices with respect to the current bag. At a given dynamic programming step, we will assume that the current solution resolves all vertex pairs in \(G_i\). Such a vertex pair may be resolved by a vertex from \(G - G_i\), or by a vertex in \(G_i\) itself.

Any bag \(X_i\) of the tree decomposition whose node \(i\) lies on a path between two join nodes in \(T\), forms a separator of \(G\): there are no edges between the vertices of \(G_i - X_i\) and \(G - G_i\). For a vertex \(v\) not in \(X_i\), we consider its distance-vector to the vertices of \(X_i\); the distance-vectors induce an equivalence relation on the vertices of \(G - X_i\), whose classes we call \(X_i\)-classes. Consider the two subgraphs \(G_i\) and \(G - G_i\). Any two solution vertices \(x, y\) from \(G - G_i\) that are in the same \(X_i\)-class, resolve the exact same pairs of vertices from \(G_i\). Thus, for this purpose, it is irrelevant whether \(x\) or \(y\) will be in a resolving set, and it is sufficient to know that a vertex of their \(X_i\)-class will eventually be chosen. In this way, one can check whether a vertex pair from \(G_i\) is resolved by a solution vertex of \(G - G_i\).

The same idea is used to “remember” the previously computed solution: it is sufficient to remember the \(X_i\)-classes of the vertices in the previously computed resolving set, rather than the vertices themselves.
It is slightly more delicate to make sure that vertex pairs in $G_i$ are resolved in the case where such a pair is resolved by a vertex in $G_i$. Indeed, this must be ensured, in particular when processing a join node $i$, for vertex pairs belonging to bags in the two sub-trees corresponding to the children $i_1, i_2$ of $i$. Such pairs may be resolved by four types of solution vertices: from $G - G_i$, $X_i, G_{i_1} - X_i$, or $G_{i_2} - X_i$. To ensure this, the dynamic programming scheme makes sure that, at each step, for any possible pair $C_1, C_2$ of $X_i$-classes, all vertex pairs $(u, v)$ consisting of a vertex $u$ of $G_i$ with class $C_1$ and a vertex $v$ of $G - G_i$ with class $C_2$ are resolved. The crucial step here is that when a new vertex $v$ is introduced (i.e., added to a bag $X_i$ to form $X_{i'}$), depending on its $X_i$-class, it must be made sure that it is resolved from all other vertices depending on their $X_i$-classes, as described above. To ensure that $v$ is distinguished from all other vertices of $G_i$, we keep track of vertex pairs of $G_i \times (G - G_i)$ that are already resolved by the partial solution, and enforce that, when processing bag $X_{i'}$, for every vertex $x$ of $G_i$, the pair $(x, v)$ is already resolved. As $v$ belongs to the new bag $X_{i'}$, we know its distances to all resolving vertices (indeed, $X_{i'}$-classes of solution vertices can be computed from their $X_i$-classes), and thus, the information can be updated accurately.

For a bag $X_i$ and a vertex $v$ not in $X_i$, the number of possible distance vectors to the vertices of $X_i$ is at most $\text{diam}(G)^{|X_i|}$. Thus, a solution for bag $X_i$ will consist of: (i) the subset of vertices of $X_i$ selected in the solution; (ii) a subset of the $\text{diam}(G)^{|X_i|}$ possible vectors to denote the $X_i$-classes from which the currently computed solution (for $G_i$) contains at least one vertex in the resolving set; (iii) a subset of the $\text{diam}(G)^{|X_i|}$ possible vectors denoting the $X_i$-classes from which the future solution needs at least one vertex of $G - G_i$ in the resolving set; (iv) a subset of the $\text{diam}(G)^{|X_i|} \times \text{diam}(G)^{|X_i|}$ possible pairs of vectors representing the $X_i$-classes of the pairs of vertices in $G_i \times (G - G_i)$ that are already resolved by the partial solution.

### 3 Geodetic Set: Lower Bound Regarding Diameter plus Treewidth

The aim of this section is to prove the following theorem.

**Theorem 3.** Unless the ETH fails, GEODETIC SET does not admit an algorithm running in time $2^{f(\text{diam})^{o(1)}} \cdot n^{O(1)}$ for any computable function $f : \mathbb{N} \to \mathbb{N}$.

Here, we present a different reduction from 3-PARTITIONED-3-SAT to GEODETIC SET. The reduction takes as input an instance $\psi$ of 3-PARTITIONED-3-SAT on $3n$ variables and returns $(G, k)$ as an instance of GEODETIC SET such that $\text{tw}(G) = \mathcal{O}(\log(n))$ and $\text{diam}(G) = \mathcal{O}(1)$. We rely on the tool of set representation from Section 2.2, that, for convenience, we reintroduce in the context of GEODETIC SET in the next subsection.

#### 3.1 Preliminary Tool: Set Representation

For a positive integer $p$, define $F_p$ as the collection of subsets of $[2p]$ that contains exactly $p$ integers. We critically use the fact that no set in $F_p$ is contained in any other set in $F_p$ (such a collection of sets is called a Sperner family). Let $\ell$ be a positive integer such that $\ell \leq \binom{2p}{p}$. We define $\text{set-rep} : [\ell] \to F_p$ as a one-to-one function by arbitrarily assigning a set in $F_p$ to an integer in $[\ell]$. By the asymptotic estimation of the central binomial coefficient, $\binom{2p}{p} \sim \frac{4^p}{\sqrt{\pi p}}$ [36]. To get the upper bound of $p$, we scale down the asymptotic function and have $\ell \leq \frac{4^p}{\sqrt{\pi p}} = 2^\varphi$. Thus, $p = \mathcal{O}(\log \ell)$.

We will apply the existence of such a function in the context of GEODETIC SET. Suppose we have a “large” collection of vertices, say $A = \{a_1, a_2, \ldots, a_\ell\}$, and a “large” collection of vertices $C = \{c_1, c_2, \ldots, c_m\}$. Moreover, we are given a function $\phi : [m] \to [\ell]$. The basic
We rename the variables in $\psi$ of 3-Partitioned-3-SAT, with $X^\alpha$, $X^\beta$, $X^\gamma$ the partition of the variable set. From $\psi$, we construct the graph $G$ as follows. We describe the construction of $X^\alpha$, with the constructions for $X^\beta$ and $X^\gamma$ being analogous. See Figure 5 for an illustration. We rename the variables in $X^\alpha$ to $x_i^\alpha$ for $i \in [n]$.

- For every variable $x_i^\alpha$, we add the vertices $t_{2i}^\alpha$ and $f_{2i-1}^\alpha$. Formally, $A^\alpha = \{t_{2i}^\alpha, f_{2i-1}^\alpha \mid i \in [n]\}$, and hence, $|A^\alpha| = 2n$.

- For every variable $x_i^\alpha$, we add four vertices: $x_i^{\alpha,\circ}, x_i^{\alpha,\triangleleft}, x_i^{\alpha,\triangleright}, x_i^{\alpha,\circ,*}$. We make $x_i^{\alpha,\circ}$ and $x_i^{\alpha,\triangleright}$ adjacent to both $t_{2i}^\alpha$ and $f_{2i-1}^\alpha$. We make $x_i^{\alpha,\circ,*}$ adjacent to both $x_i^{\alpha,\circ}$ and $x_i^{\alpha,\triangleright}$. We make $x_i^{\alpha,\circ,*}$ adjacent to $x_i^{\alpha,\triangleright}$.

- We add the vertices $y_1, y_2, z_1, z_2$. We make $y_1$ and $y_2$ adjacent to every vertex of $A^\alpha$. We make $y_i$ adjacent to $z_i$ for $i \in \{1, 2\}$. Note that $y_1, y_2, z_1, z_2$ are common to $X^\beta$ and $X^\gamma$.

- We add the vertex $g_1$ and make it adjacent to $y_1, y_2$, and $x_i^{\alpha,\circ,*}$ for each $i \in [n]$. Note that $g_1$ is common to $X^\beta$ and $X^\gamma$. We add edges between $g_1$ and every vertex of $A^\alpha$.

Figure 5: Overview of the reduction. We only draw $A^\alpha$ and $V^\alpha$ here, as $A^\beta$, $A^\gamma$, $V^\beta$, and $V^\gamma$ are similar. The yellow lines joining $g_1$, $g_2$, $y_1$, and $y_2$ to sets indicate that the corresponding vertex is adjacent to all the vertices of the corresponding set. Suppose that $f_{2i-1}^\alpha$ and $t_{2i}^\alpha$ are in the geodetic set and $z_i$ appears in the clause $c_q$. The thick green path is a shortest path between $t_{2i}^\alpha$ and $c_q^b$ which does not cover $c_q$. The thick violet path plus the edge $(c_q^a, c_q^b)$ is a shortest path between $f_{2i-1}^\alpha$ and $c_q^b$ covering $c_q$.

The idea is to design gadgets such that $c_q$ is only covered by the shortest path from $a_{\phi(q)} \in A$ to $c_q^b$ ($c_q^b$ is forced to be chosen in the geodetic set) for any $q \in [m]$, while keeping the treewidth of this part of the graph of order $O(\log(|A|))$. To do so, we create a “small” intermediate set $V$ (of size $O(\log(|A|))$) through which will go the shortest paths between vertices in $A$ and $C$, and we connect $a_i$ to the vertices of $V$ corresponding to the bit-representation of $\text{set-rep}(i)$, and $c_q$ (with $i = \phi(q)$) to all the other vertices of $V$. In this way, the construction will ensure that $c_q$ is covered by a shortest path between $a_{\phi(q)}$ and $c_q^b$, but is not covered by any other shortest path between a vertex of $A$ and a vertex of $C$. We give the details in the following subsection.

3.2 Reduction

Consider an instance $\psi$ of 3-Partitioned-3-SAT, with $X^\alpha, X^\beta, X^\gamma$ the partition of the variable set. From $\psi$, we construct the graph $G$ as follows. We describe the construction of $X^\alpha$, with the constructions for $X^\beta$ and $X^\gamma$ being analogous. See Figure 5 for an illustration. We rename the variables in $X^\alpha$ to $x_i^\alpha$ for $i \in [n]$.

- For every variable $x_i^\alpha$, we add the vertices $t_{2i}^\alpha$ and $f_{2i-1}^\alpha$. Formally, $A^\alpha = \{t_{2i}^\alpha, f_{2i-1}^\alpha \mid i \in [n]\}$, and hence, $|A^\alpha| = 2n$.

- For every variable $x_i^\alpha$, we add four vertices: $x_i^{\alpha,\circ}, x_i^{\alpha,\circ,*}, x_i^{\alpha,\triangleright}, x_i^{\alpha,\circ,*}$. We make $x_i^{\alpha,\circ}$ and $x_i^{\alpha,\triangleright}$ adjacent to both $t_{2i}^\alpha$ and $f_{2i-1}^\alpha$. We make $x_i^{\alpha,\circ,*}$ adjacent to both $x_i^{\alpha,\circ}$ and $x_i^{\alpha,\triangleright}$. We make $x_i^{\alpha,\circ,*}$ adjacent to $x_i^{\alpha,\triangleright}$.

- We add the vertices $y_1, y_2, z_1, z_2$. We make $y_1$ and $y_2$ adjacent to every vertex of $A^\alpha$. We make $y_i$ adjacent to $z_i$ for $i \in \{1, 2\}$. Note that $y_1, y_2, z_1, z_2$ are common to $X^\beta$ and $X^\gamma$.

- We add the vertex $g_1$ and make it adjacent to $y_1, y_2$, and $x_i^{\alpha,\circ,*}$ for each $i \in [n]$. Note that $g_1$ is common to $X^\beta$ and $X^\gamma$. We add edges between $g_1$ and every vertex of $A^\alpha$.
Let $p$ be the smallest positive integer such that $2n \leq (2^p)^{2p}$. In particular, $p = \Theta(\log n)$.

We add a validation portal, a clique on $2p$ vertices, denoted by $V^\alpha = \{v_1^\alpha, v_2^\alpha, \ldots, v_{2p}^\alpha\}$.

For each $\delta \in \{\alpha, \beta, \gamma\}$, we add edges between $g_1$ and every vertex of $V^\delta$.

For every clause $C_q$ in $\psi$, we introduce three vertices: $c_q, c_q^\alpha, c_q^\gamma$. We add the edges $(c_q, c_q^\alpha)$ and $(c_q^\alpha, c_q^\gamma)$.

Define set-rep : $[2n] \mapsto \mathcal{F}_p$ as an arbitrary injective function, where $\mathcal{F}_p$ is the Sperner family (and $p$ is as defined two items above). Add the edge $(t_{2i}^q, v_{2p}^\alpha)$ for every $p' \in \text{set-rep}(2i)$ and the edge $(f_{2i-1}^\alpha, v_{2p}^\gamma)$ for every $p' \in \text{set-rep}(2i - 1)$. If the variable $x_i^\delta$ appears positively in the clause $C_q$, then we add the edges $(c_q, v_{2p}^\alpha)$ and $(c_q^\alpha, v_{2p}^\gamma)$ for every $p' \in [2p] \setminus \text{set-rep}(2i - 1)$.

If the variable $x_i^\delta$ appears negatively in the clause $C_q$, then we add the edges $(c_q, v_{2p}^\alpha)$ and $(c_q^\alpha, v_{2p}^\gamma)$ for every $p' \in [2p] \setminus \text{set-rep}(2i - 1)$.

Add a vertex $g_2$ and make $g_2$ adjacent to every vertex of $A^\alpha$ and every vertex of $\{c_q : q \in [m]\}$. Note that $g_2$ is common to $X^\beta$ and $X^\gamma$.

Add a vertex $g_3$ and make it adjacent to every vertex of $\{c_q^\alpha : q \in [m]\}$. Note that $g_3$ and the vertices of $\{c_q, c_q^\alpha, c_q^\gamma : q \in [m]\}$ are common to $X^\beta$ and $X^\gamma$.

This concludes the construction of $G$. The reduction returns $(G, k)$ as an instance of GEODETIC SET where $k = 6n + m + 2$.

### 3.3 Correctness of the Reduction

Suppose, given an instance $\psi$ of 3-PARTITIONED-3-SAT, that the reduction above returns $(G, k)$ as an instance of GEODETIC SET.

**Lemma 4.** If $\psi$ is a satisfiable 3-PARTITIONED-3-SAT formula, then $G$ admits a geodetic set of size $k$.

**Proof.** Suppose that $\pi : X^\alpha \cup X^\beta \cup X^\gamma \mapsto \{\text{True}, \text{False}\}$ is a satisfying assignment for $\psi$. We construct a geodetic set $S$ of size $k$ for $G$ using this assignment.

For every $\delta \in \{\alpha, \beta, \gamma\}$ and $i \in [n]$, if $\pi(x_i^\delta) = \text{True}$, then let $t_{2i}^q \in S$, and otherwise, $f_{2i-1}^\delta \in S$. We also put $z_1, z_2, x_i^\delta, x_i^\delta^*$, and $c_q^\delta$ into $S$ for all $i \in [n], \delta \in \{\alpha, \beta, \gamma\}$, and $q \in [m]$. Note that $|S| = k$.

Now, we show that $S$ is indeed a geodetic set of $G$. First, $g_1, g_2, z_1, z_2, g_1$, and all the vertices of $A^\alpha, A^\beta, A^\gamma$ are covered by a shortest path between $z_1$ and $z_2$. Then, for each $\delta \in \{\alpha, \beta, \gamma\}$ and $i \in [n]$, $x_i^\delta, x_i^\delta^*$, $x_i^\delta$, and $x_i^\delta^*$ are covered by a shortest path between $S \cap \{t_{2i}^q, f_{2i-1}^\delta\}$ and $x_i^\delta^*$. The vertex $g_3$ is covered by any shortest path between $c_q^\delta$ and $c_q^\delta$, where $C_q$ and $C_q'$ are two clauses of $\psi$. Suppose that $\pi(x_i^\delta)$, for some $i \in [n]$ and $\delta \in \{\alpha, \beta, \gamma\}$, satisfies some clause $C_q$. By our construction, if $x_i^\delta$ appears positively (negatively, respectively) in $C_q$, then $t_{2i}^q$ ($f_{2i-1}^\delta$, respectively) and $c_q^\delta$ are at distance four since $t_{2i}^q$ ($f_{2i-1}^\delta$, respectively) and $c_q^\delta$ have no common neighbor in $V^\delta$. Moreover, there is a shortest path from $t_{2i}^q$ ($f_{2i-1}^\delta$, respectively) to $c_q^\delta$ of length four, covering $g_2, c_q, c_q^\delta$, and $c_q^\delta$; there is also a shortest path from $t_{2i}^q$ ($f_{2i-1}^\delta$, respectively) to $c_q^\delta$ of length four, covering $c_q^\delta$, $v_k$, $v_q$, and $c_q^\delta$, where $v_k^\delta \in V^\delta$ is a vertex adjacent to $t_{2i}^q$ ($f_{2i-1}^\delta$, respectively) and $v_k^\delta$ is any vertex of $V^\delta$ that is not adjacent to $t_{2i}^q$ ($f_{2i-1}^\delta$, respectively). Thus, every vertex of $V^\delta$ for $\delta \in \{\alpha, \beta, \gamma\}$ is covered by a shortest path between two vertices of $S$. Since every clause of $\psi$ is satisfied by $\pi$, it follows that every vertex of $\{c_q, c_q^\alpha, c_q^\gamma : q \in [m]\}$ is covered by a shortest path between two vertices of $S$. As a result, $S$ is a geodetic set of $G$. \hfill \blacksquare
Proof. Suppose that $G$ has a geodetic set $S$ of size at most $k$. Since they all have degree 1, $z_1, z_2$, $x_i^{q, a}$, and $c_q$ for all $i \in [n]$, $\delta \in \{\alpha, \beta, \gamma\}$, and $q \in [m]$ must be in any geodetic set $S$ of $G$.

Claim 6. For each $i \in [n]$ and $\delta \in \{\alpha, \beta, \gamma\}$, exactly one of $t_{2i}^\delta$ and $f_{2i-1}^\delta$ must be in $S$.

Proof. Since $S$ is a geodetic set, for each $i \in [n]$ and $\delta \in \{\alpha, \beta, \gamma\}$, $x_i^{\delta, a}$ and $x_i^{\delta, b}$ must be covered by shortest paths between two vertices of $S$. If $t_{2i}^\delta \in S (f_{2i-1}^\delta \in S$, respectively), $x_i^{\delta, a}$ and $x_i^{\delta, b}$ are covered by shortest paths between $t_{2i}^\delta \in S$ (respectively), and $x_i^{\delta, *}$.

Suppose that, for some $i' \in [n]$ and $\delta' \in \{\alpha, \beta, \gamma\}$, neither of $t_{2i'}^\delta$ and $f_{2i'-1}^\delta$ is in $S$. Moreover, if neither of $x_i^{\delta, a}$ and $x_i^{\delta, b}$ is in $S$, then, due to the edges incident with $g_1$, no vertices in $S$ have a shortest path containing any of these two vertices. Similarly, if only one of $x_i^{\delta, a}$ and $x_i^{\delta, b}$ is in $S$, then the other is not covered by $S$. Thus, if neither of $t_{2i'}^\delta$ and $f_{2i'-1}^\delta$ is in $S$, then both $x_i^{\delta, a}$ and $x_i^{\delta, b}$ must be in $S$. Since $k - |\{z_1, z_2\} \cup \{x_i^{\delta, q} : i \in [n], \delta \in \{\alpha, \beta, \gamma\}\} \cup \{c_q : q \in [m]\} = 3n$, we conclude that exactly one of $t_{2i}^\delta$ and $f_{2i-1}^\delta$ must be in $S$ for each $i \in [n]$ and $\delta \in \{\alpha, \beta, \gamma\}$. 

By Claim 6 and earlier arguments, we now have that $|S| = k$.

Claim 7. For each $q \in [m]$, the vertex $c_q$ is covered either by a shortest path between $c_q^a$ and $t_{2i}^\delta$, where the variable $x_i^q$ appears positively in the clause $C_q$, or by a shortest path between $c_q^b$ and $f_{2i-1}^\delta$, where the variable $x_i^q$ appears negatively in the clause $C_q$. Moreover, $c_q$ is covered by no other type of shortest path between two vertices in $S$.

Proof. By the construction of $G$, if the variable $x_i^q$ appears positively in the clause $C_q$, then there is a shortest path from $t_{2i}^\delta$ to $c_q^a$ of length four covering $g_2, c_q, c_q^a$, and $c_q^\delta$. If the variable $x_i^q$ appears negatively in the clause $C_q$, then there is a shortest path from $f_{2i-1}^\delta$ to $c_q^b$ of length four covering $g_2, c_q, c_q^b$, and $c_q^\delta$.

Next, we show that $c_q$ is not covered by any shortest path between any other two vertices of $S$. We can check that $c_q$ is not covered by any of the shortest paths between $z_1$ and $z_2$, between $z_j (j \in \{1, 2\})$ and $x_i^{\delta, *}$ ($i \in [n], \delta \in \{\alpha, \beta, \gamma\}$), and between $z_j (j \in \{1, 2\})$ and $S \cap \{t_{2i}, f_{2i-1}^\delta (i \in [n], \delta \in \{\alpha, \beta, \gamma\}\})$. Note that any shortest path from $z_j (j \in \{1, 2\})$ to $c_q^a (q \in [m])$ is of length five, covering $g_1$, some vertex of $A^\delta (\delta \in \{\alpha, \beta, \gamma\}$), some vertex of $V^\delta, c_q^a$, and $c_q^\delta$.

We can check that $c_q$ is not covered by any of the shortest paths between $x_i^{\delta, a}$ and $x_i^{\delta, b}$ ($i, i' \in [n], \delta, \delta' \in \{\alpha, \beta, \gamma\}$), and between $x_i^{\delta, *}$ and $S \cap \{t_{2i}, f_{2i-1}^\delta (i, i' \in [n], \delta, \delta' \in \{\alpha, \beta, \gamma\}\}$. Note that any shortest path from $x_i^{\delta, a} (i \in [n], \delta \in \{\alpha, \beta, \gamma\})$ to $c_q (q \in [m])$ is of length five, covering $x_i^{\delta, a}, g_1$, some vertex of $V^\delta, c_q^a$, and $c_q^\delta$.

By the case analysis above, the claim is true.

By Claim 6, exactly one vertex of $t_{2i}^\delta$ and $f_{2i-1}^\delta$ belongs to $S$ for each $i \in [n]$ and $\delta \in \{\alpha, \beta, \gamma\}$. We define an assignment $\pi$ to the variables of $\psi$ as follows. For each $i \in [n]$ and $\delta \in \{\alpha, \beta, \gamma\}$, if $t_{2i}^\delta \in S$, then $\pi(x_i^\delta) = \text{True}$. Otherwise, $\pi(x_i^\delta) = \text{False}$. Since $S$ is a geodetic
set for \( G \), every vertex \( c_q (q \in [m]) \) is covered by a shortest path between two vertices of \( S \). By Claim 7, every vertex \( c_q (q \in [m]) \) is covered by a shortest path between \( S \cap \{ t_{2i}, f_{2i-1} \} \) and \( c_q \), where the variable \( x_i^q \) appears in the clause \( C_q \). It follows that every clause \( C_q \) is satisfied by \( \pi(x_i^q) \). As a result, \( \psi \) is a satisfiable 3-PARTITIONED-3-SAT formula.

Proof of Theorem 3. First, it is not hard to check that the diameter of \( G \) is at most 5. Then, let \( X = V^\alpha \cup V^\beta \cup V^\gamma \cup \{ g_1, g_2, g_3, y_1, y_2 \} \). We can check that every component of \( G \setminus X \) has at most six vertices and \( |X| = O(\log n) \). Thus, the treewidth \( tw(G) \) – in fact, even the treebreadth \( td(G) \) – of \( G \) is bounded by \( O(\log n) \). By the description of the reduction, it takes polynomial time to compute the reduced instance. Hence, if there is an algorithm for GEODETIC SET that runs in time \( 2^{f(\text{diam})}o(tw) \) (or \( 2^{f(\text{diam})}o(td) \)), then there is an algorithm running in time \( 2^{o(n)} \) for 3-PARTITIONED-3-SAT, which contradicts the ETH.

4 Conclusion

We have shown (under the ETH) that three natural metric-based graph problems, METRIC DIMENSION, GEODETIC SET, and STRONG METRIC DIMENSION, exhibit tight (double-) exponential running times for the standard structural parameterizations by treewidth and vertex cover number. This includes tight double-exponential running times for treewidth plus diameter (METRIC DIMENSION and GEODETIC SET) and for vertex cover (STRONG METRIC DIMENSION).

Such tight double-exponential running times for FPT structural parameterizations of graph problems had previously been observed only for counting problems and problems complete for classes above \( \text{NP} \). Thus, surprisingly, our results show that some natural problems can be in \( \text{NP} \) and still exhibit such a behavior.

It would be interesting to see whether this phenomenon holds for other graph problems in \( \text{NP} \), and for other structural parameterizations. Perhaps one can determine certain properties shared by these metric-based graph problems, that imply such running times, with the goal of generalizing our approach to a broader class of problems. In particular, concerning the general versatile technique that we designed to obtain the double-exponential lower bounds, it would be intriguing to see for which other problems in \( \text{NP} \) our technique works.

In fact, after this paper appeared online, our technique was successfully applied to an \( \text{NP} \)-complete problem in machine learning [11] (for \( \text{vc} \)) as well as \( \text{NP} \)-complete identification problems [10] (for \( \text{tw} \)).

References


Problems in NP Can Admit Double-Exponential Lower Bounds


