A polynomial kernel for Distance-Hereditary Vertex Deletion

Eun Jung Kim[1] and O-joung Kwon[2] ⋆

[1] CNRS-Université Paris-Dauphine, Place du Marechal de Lattre de Tassigny, 75775 Paris cedex 16, France
[2] Logic and Semantics, TU Berlin, Berlin, Germany
eunjungkim78@gmail.com, ojoungkwon@gmail.com

Abstract. A graph is *distance-hereditary* if for any pair of vertices, their distance in every connected induced subgraph containing both vertices is the same as their distance in the original graph. The Distance-Hereditary Vertex Deletion problem asks, given a graph G on n vertices and an integer k, whether there is a set S of at most k vertices in G such that $G - S$ is distance-hereditary. This problem is important due to its connection to the graph parameter rank-width [19]; distance-hereditary graphs are exactly the graphs of rank-width at most 1. Eiben, Ganian, and Kwon (MFCS' 16) proved that Distance-Hereditary Vertex Deletion can be solved in time $2^{\mathcal{O}(k)}n^{\mathcal{O}(1)}$, and asked whether it admits a polynomial kernelization. We show that this problem admits a polynomial kernel, answering this question positively. For this, we use a similar idea for obtaining an approximate solution for Chordal Vertex Deletion due to Jansen and Pilipczuk (SODA' 17) to obtain an approximate solution with $\mathcal{O}(k^3 \log n)$ vertices when the problem is a Yes-instance, and we exploit the structure of split decompositions of distance-hereditary graphs to reduce the total size.

1 Introduction

A graph is *distance-hereditary* if for every connected induced subgraph H and two vertices u and v in H, the distance between u and v in H is the same as their distance in G. A vertex subset X of a graph G is a *distance-hereditary modulator*, or a *DH-modulator* in short, if $G - X$ is a distance-hereditary graph. We study the problem Distance-Hereditary Vertex Deletion (DH Vertex Deletion) which asks, given a graph G and an integer k, whether G contains a DH-modulator of size at most k.

The graph modification problems, in which we want to transform a graph to satisfy a certain property with as few graph modifications as possible, have been extensively studied. For instance, the Vertex Cover and Feedback Vertex Set problems are graph modification problems where the target graphs

⋆ O. Kwon is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC consolidator grant DISTRUCT, agreement No. 648527).

F. Ellen et al. (Eds.): WADS 2017, LNCS 10389, pp. 509–520, 2017.
DOI: 10.1007/978-3-319-62127-2_43

are edgeless graphs and forests, respectively. By the classic result of Lewis and Yannakakis [18], it is known that for all non-trivial hereditary properties that can be tested in polynomial time, the corresponding vertex deletion problems are NP-complete. Hence, the research effort has been directed toward designing algorithms such as approximation and parameterized algorithms.

When the target graph class $\mathcal{C}$ admits efficient recognition algorithms for some NP-hard problems, the graph modification problem related to such a class attracts more attention. Vertex deletion problems to classes of graphs of constant tree-width or constant tree-depth have attracted much attention in this context. Tree-width w Vertex Deletion is proved to admit an FPT algorithm running in time $2^{\mathcal{O}(k)}n^{\mathcal{O}(1)}$ and a kernel with $\mathcal{O}(k^{g(w)})$ vertices for some function g [11,17]. Also, it was shown that Tree-depth w Vertex Deletion admits uniformly polynomial kernels with $\mathcal{O}(k^6)$ vertices, for every fixed w [12]. All these problems are categorized as vertex deletion problems for $\mathcal{F}$-minor free graphs in a general setting, when the set $\mathcal{F}$ contains at least one planar graph. However, $\mathcal{F}$-minor free graphs capture only sparse graphs in a sense that the number of edges of such a graph is bounded by a linear function on the number of its vertices. Thus these problems are not useful when dealing with very dense graphs.

Rank-width [19] and *clique-width* [5] are graph width parameters introduced for extending graph classes of bounded tree-width. Graphs of bounded rank-width represent graphs that can be recursively decomposed along vertex partitions (X, Y) where the number of neighborhood types between X and Y are small. Thus, graphs of constant rank-width may contain dense graphs; for instance, all complete graphs have rank-width at most 1. Courcelle, Makowski, and Rotics [4] proved that every MSO_1-expressible problem can be solved in polynomial time on graphs of bounded rank-width.

Motivated from Tree-width w Vertex Deletion, Eiben, Ganian, and the second author [9] initiated study on vertex deletion problems to graphs of constant rank-width. The class of graphs of rank-width at most 1 is exactly the class of distance-hereditary graphs [19]. It was known that the vertex deletion problem for graphs of rank-width w can be solved in FPT time [16] using a meta-theorem [4]. Eiben et al. [9] devised the first elementary algorithm for this problem when $w = 1$, or equivalently DH Vertex Deletion, that runs in time $2^{\mathcal{O}(k)}n^{\mathcal{O}(1)}$. Furthermore, they discussed that a DH-modulator of the size k can be used to obtain a $2^{\mathcal{O}(k)}n^{\mathcal{O}(1)}$-time algorithm for problems such as Independent Set, Vertex Cover, and 3-Coloring.

However, until now, it was not known whether DH Vertex Deletion admits a polynomial kernel or not. A *kernelization* of a parameterized graph problem Π is a polynomial-time algorithm which, given an instance (G, k) of Π, outputs an equivalent instance (G', k') of Π with $|V(G')| + k' \leq h(k)$ for some computable function h. The resulting instance (G', k') of a kernelization is called a *kernel*, and in particular, when h is a polynomial function, Π is said to admit a *polynomial kernel*.

Our Contribution and Approach. Our main result is the following.

Theorem 1. DH Vertex Deletion *admits a polynomial kernel.*

We introduce in Section 3 an approximate DH-modulator with $\mathcal{O}(k^3 \log n)$ vertices if the given instance is a Yes-instance. An important observation here is that a distance-hereditary graph contains a complete bipartite subgraph (not necessarily induced) which is a balanced separator. Thus, if G admits a small DH-modulator, then there is balanced vertex separator $X \uplus K$ where X is small and K induces a complete bipartite subgraph. By recursively extracting such separators using an approximation algorithm for finding a balanced vertex separator [10], we will decompose the given graph into $D \uplus K_1 \uplus \cdots \uplus K_\ell \uplus X$, where $\ell = \mathcal{O}(k \log n)$, D is distance-hereditary, each K_i is a complete bipartite subgraph, $|X| = \mathcal{O}(k^2 \sqrt{\log k} \log n)$. In the next step, we argue that if a graph H is a disjoint union of a distance-hereditary graph and a complete bipartite graph and (H, k) is a Yes-instance and satisfies a certain property, then in polynomial time, one can construct a DH-modulator of size $\mathcal{O}(k^2)$ for H (Proposition 2). Using this sub-algorithm ℓ times, we construct an approximate DH-modulator with $\mathcal{O}(k^3 \log n)$ vertices. This part follows a vein similar to the approach of Jansen and Pilipczuk [15] for Chordal Vertex Deletion. Given a DH-modulator S of size $\mathcal{O}(k^3 \log n)$, we can obtain a new DH-modulator S' of size $\mathcal{O}(k^5 \log n)$ such that for every $v \in S'$, $G[(V(G) \setminus S') \cup \{v\}]$ is also distance-hereditary by adding $\mathcal{O}(k^2)$ vertices per each vertex in S. Such a DH-modulator is called a *good DH-modulator* and the details will be explained in Section 4.

The remaining part of the paper is contributed to reduce the number of vertices in $G - S'$. Two vertices v and w are twins if they have the same neighbors outside $\{v, w\}$. In Section 5, we present a reduction rule that bounds the size of each set of pairwise twins in $G - S'$. We give, in Section 6, a reduction rule that bounds the number of components of $G - S'$. Lastly in Section 7, we reduce the size of each component of $G - S'$ having at least 2 vertices. For the last part, we use split decompositions of distance-hereditary graphs. Briefly, a split decomposition displays a tree-like structure of a distance-hereditary graph in the form of a decomposition tree with bags for each nodes, such that each bag consists of a maximal set of pairwise twins (possibly with an extra vertex) in $G - S'$. We will provide a rule that bounds the number of bags in the decomposition tree, which results in bounding the size of each component.

2 Preliminaries

We follow [8] for basic graph terminology. A graph is *trivial* if it consists of a single vertex, and *non-trivial* otherwise. For two sets $A, B \subseteq V(G)$, we say A is *complete* to B if for every $v \in A$ and $w \in B$, v is adjacent to w. Two vertices v and w of a graph G are *twins* if they have the same neighbors in $G - \{v, w\}$. A vertex partition (A, B) of G is *split* if $N_G(A)$ is complete to $N_G(B)$.

A graph H is a *biclique* if there is a bipartition of $V(H)$ into non-empty sets $A \uplus B$ such that any two vertices $a \in A$ and $b \in B$ is adjacent. Notice that there may be edges among the vertices of A or B. For $K \subseteq V(G)$, we say that K is a biclique of G if $G[K]$ is a biclique. For a connected graph G, a vertex subset S

of G is called a *balanced vertex separator* of G if every component of $G - S$ has at most $\frac{2}{3}|V(G)|$ vertices. We allow $V(G)$ to be a balanced vertex separator of G. For a vertex subset S of G, a path is called an *S-path* if its end vertices are in S and all other internal vertices are in $V(G) \setminus S$.

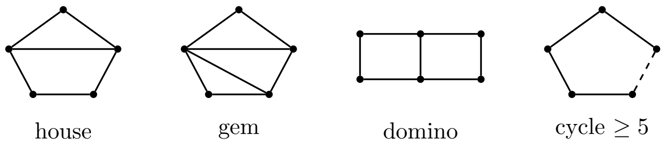

Fig. 1. The induced subgraph obstructions for distance-hereditary graphs.

A graph is called a *DH obstruction* if it is isomorphic to a gem, a house, a domino or an induced cycle of length at least 5, that are depicted in Figure 1. A DH obstruction is *small* if it has at most 6 vertices. Bandelt and Mulder [2] proved that a graph is distance-hereditary if and only if it has no induced subgraph isomorphic to one of DH obstructions. A DH-modulator S is *good* if $G[(V(G) \setminus S) \cup \{v\}]$ is distance-hereditary for every $v \in S$.

3 Approximation algorithm

We present a polynomial-time algorithm which constructs an approximate DH-modulator of G whenever (G, k) is a Yes-instance.

Theorem 2. *There is a polynomial-time algorithm which, given a graph G and a positive integer k, either correctly reports that (G, k) is a* No*-instance to* DH Vertex Deletion*, or returns a DH-modulator of size $O(k^3 \log n)$.*

If G contains $k+1$ vertex-disjoint copies of small DH obstructions, then (G, k) is a No-instance. We may assume a maximal packing of small DH obstructions has cardinality at most k. Since a maximal packing consists of at most $6k$ vertices, it is sufficient to prove Theorem 2 when G has no small DH obstruction.

We prove the following two propositions, implying Theorem 2 together.

Proposition 1. *There is a polynomial-time algorithm which, given an instance (G, k), either computes a decomposition $V(G) = D \uplus K_\ell \uplus \cdots \uplus K_1 \uplus X$ such that $G[D]$ is distance-hereditary, each K_i is a biclique, $|X| = O(k^2\sqrt{\log k}\log n)$ and $\ell = O(k \log n)$, or correctly reports that (G, k) is a* No*-instance.*

When G does not contain any small DH obstructions, a linear program of DH Vertex Deletion for G can be formulated as follows, where $x_v \geq 0$ for each $v \in V(G)$:

$$\min \sum_{v \in V(G)} x_v \ s.t \quad \sum_{v \in V(H)} x_v \geq 1 \qquad \forall \text{ induced cycle } H \text{ of length at least } 7$$

A mapping $x = (x_v)_{v \in V(G)}$ from $V(G)$ to $\mathbb{R}$ is a *feasible fractional solution* to DH Vertex Deletion for G if it is feasible to the above linear program for G. For a subgraph H of G, we write $x(H) := \sum_{v \in V(H)} x_v$ and $|x| := x(G)$.

Proposition 2. *Let (G, k) be an instance to* DH VERTEX DELETION *such that G has no small DH obstructions, $V(G) = D \uplus K$ where $G[D]$ is distance-hereditary, and K is a biclique. Let x^* be a feasible fractional solution to* DH VERTEX DELETION *for G such that $x^*_v < \frac{1}{20}$, $\forall v \in V(G)$. Given such G and x^*, one can in polynomial time find a DH-modulator X with $O(|x^*|^2)$ vertices.*

We first explain Proposition 1. First, we obtain an $\mathcal{O}(n^3)$ bound on the number of maximal bicliques in a graph having no small DH obstructions. Secondly, we prove that every connected distance-hereditary graph on at least two vertices contains a balanced vertex separator that is a biclique. Combining these results, we can show the following.

Lemma 1. *Whenever (G, k) is a* YES*-instance and G has no small DH obstructions, one can in polynomial time find a balanced vertex separator $K \uplus X$ where K is a biclique or an empty set and $|X| = \mathcal{O}(k\sqrt{\log k})$.*

Proof (Sketch of proof). Over all maximal bicliques K of G, we apply the $\mathcal{O}(\sqrt{\log OPT})$-approximation algorithm for finding a balanced vertex separator in $G - K$, due to Feige et al [10]. One can observe that since (G, k) is a YES-instance, there is some set X of size at most k and a balanced vertex separator of K' of $G - X$ that is a biclique. Thus, a maximal biclique of G containing this K' is detected in the algorithm, and the approximation algorithm provides a set X' of size $\mathcal{O}(k\sqrt{\log k})$ where $K' \uplus X'$ is a balanced vertex separator. □

We set $G_1 := G$, $K_0 = X_0 = \emptyset$, and at i-th recursive step, we apply Lemma 1 to a connected component G_i of $G - \bigcup_{j<i}(K_j \uplus X_j)$ which is not distance-hereditary and obtain a balanced vertex separator $K_i \uplus X_i$ of G_i. In the end, we obtain a decomposition $V(G) = D \uplus K_\ell \uplus \cdots \uplus K_1 \uplus X_\ell \uplus \cdots \uplus X_1$, where $G[D]$ is distance-hereditary, each K_i is a biclique or an empty set, and $|X_i| = O(k\sqrt{\log k})$. Since we only apply Lemma 1 to a component that is not distance-hereditary, if (G, k) is a YES-instance, then the size-k-modulator of G intersects every such component. By representing the recursive procedure as a collection of branching trees $\mathcal{T}$, we can show that $\ell = O(k \log n)$, as the maximum length of a root-to-leaf path in $\mathcal{T}$ is $O(\log n)$.

Now, we explain Proposition 2. Suppose G has no small DH obstructions and $V(G) = D \uplus K$ where $G[D]$ is distance-hereditary, and K is a biclique with a bipartition (A, B), and x^* is a feasible fractional solution to DH VERTEX DELETION such that $x^*_v < \frac{1}{20}$ for every $v \in V(G)$. We first observe that a new vector x' where $x'_v = 0$ if $v \in K$ and $x'_v = 2x^*_v$ if $v \in D$, is again a feasible fractional solution. For this, we show that every induced cycle H of length at least 7 in G satisfies that if $G[V(H) \cap K]$ has one component, then $|V(H) \cap K| \leq 3$, and otherwise, H contains a K-path whose length is at least 3. In the former case, we have $x'(H) = x'(G[V(H) \cap D]) = 2x^*(G[V(H) \cap D]) \geq 1$ as $x^*(G[V(H) \cap K]) < 3 \cdot \frac{1}{20} < \frac{1}{2}$. In the latter case, the end vertices of the K-path P are contained in the same part of A or B, and it forms another DH obstruction H' with a vertex in the other part, where $G[V(H') \cap K]$ has one component. Thus, we have $x'(H) \geq x'(P) \geq x'(H') \geq 1$, as $x'_v = 0$ for $v \in K$.

We construct an instance $(G[D], \mathcal{T})$ of Vertex Multicut with terminal pairs $\mathcal{T} := \{(s,t) \subseteq D \times D : \mathrm{dist}_{G[D],x'}(s,t) \geq 1\}$, where $\mathrm{dist}_{G[D],x'}(s,t)$ is the minimum $x'(P)$ over all (s,t)-paths P. Notice that for every terminal pair $(s,t) \in \mathcal{T}$, and for every (s,t)-path P in $G[D]$, we have $x'(P) \geq \mathrm{dist}_{G[D],x'}(s,t) \geq 1$, meaning that x' is a feasible fractional solution to Vertex Multicut for the instance $(G[D], \mathcal{T})$. Using an approximation algorithm for Vertex Multicut by Gupta [14], we can obtain a vertex set $X \subseteq D$ of size $O(|x'|^2)$ such that $G[D \setminus X]$ contains no (s,t)-path for every terminal pair $(s,t) \in \mathcal{T}$ in polynomial time. We prove, in the appendix, that the obtained set X is a DH-modulator.

Proof (of Theorem 2). It is sufficient to prove when G has no small DH obstructions. Let x^* be an optimal fractional solution to DH Vertex Deletion for G. We may assume $|x^*| \leq k$, otherwise we report that (G,k) is a No-instance. Let $\tilde{X}$ be the set of all vertices v such that $x^*_v \geq \frac{1}{20}$. Observe that $|\tilde{X}| \leq 20k$ since otherwise, $|x^*| \geq \frac{1}{20}|\tilde{X}| > k$, a contradiction. Also x^* restricted to $V(G) \setminus \tilde{X}$ is a fractional feasible solution for $G - \tilde{X}$ such that $x^*_v < \frac{1}{20}$ for every v.

We compute a decomposition $V(G - \tilde{X}) = D \uplus \bigcup_{i=1}^{\ell} K_i \cup X$ as in Proposition 1, or correctly report (G,k) as a No-instance. Recall that $\ell = O(k \log n)$ and $|X| = O(k^2\sqrt{\log k}\log n)$. Note that $V(G - (\tilde{X} \cup X)) = D \uplus \bigcup_{i=1}^{\ell} K_i$. From $i = 1$ up to ℓ, we want to obtain a DH-modulator S_i of G_i, where $G_1 := G[D \cup K_1]$ and for $i = 2, \ldots, \ell$, G_i is the subgraph of G induced by $(V(G_{i-1}) \setminus S_{i-1}) \cup K_i$. Note that $G_i - S_i$ is distance-hereditary and K_i is a blique. Hence, we can inductively apply the algorithm of Proposition 2 and obtain a DH-modulator S_i of size at most $O(|x^*|^2)$ of G_i. Especially, $G_\ell - S_\ell$ is distance-hereditary, implying that the set defined as $S := \tilde{X} \cup X \cup \bigcup_{i=1}^{\ell} S_i$ is a DH-modulator of G. From $|x^*| \leq k$, we have $|S_i| = O(k^2)$ for each i. It follows that $|S| = O(k^3 \log n)$. □

4 Good modulator

Theorem 3. *There is a polynomial-time algorithm which, given a graph G and a positive integer k, either correctly reports that (G,k) is a* No*-instance, or returns an equivalent instance (G', k') with a good DH-modulator of size $O(k^5 \log n)$.*

Proof (Sketch of Proof). If the algorithm of Theorem 2 reports that the instance is a No-instance, then we are done. Let S be a DH-modulator of size $\mathcal{O}(k^3 \log n)$ given by Theorem 2. Let $U := \emptyset$, and for $v \in S$, let $H_v := G[(V(G) \setminus S) \cup \{v\}]$. One can in polynomial time find either $k+1$ small DH obstructions in H_v whose pairwise intersection is v, or a vertex set T_v of $V(G) \setminus S$ such that $|T_v| \leq 5k$ and $H_v - T_v$ has no small DH obstructions. In the former case, we add v to U.

Assume we obtain a vertex set T_v. Since $H_v - T_v$ has no small DH obstructions, every DH obstruction in $H_v - T_v$ is an induced cycle of length at least 7. We assert that either $H_v - T_v$ contains a vertex set X_v of size $\mathcal{O}(k^2)$ such that $H_v - (T_v \cup X_v)$ has no DH obstructions, or correctly reports that every DH-modulator of size at most k contains v.

We consider an instance $(H_v - (T_v \cup \{v\}), \mathcal{T})$ of Vertex Multicut where $\mathcal{T} := \{(s,t) : s,t \in N_{H_v - T_v}(v), \mathrm{dist}_{H_v - (T_v \cup \{v\})}(s,t) \geq 3\}$. We can show that

$X \subseteq V(H_v) \setminus (T_v \cup \{v\})$ hits all induced cycles of $H_v - T_v$ of length at least 7 if and only if X is a vertex multicut for $(H_v - (T_v \cup \{v\}), \mathcal{T})$, because the restriction of an induced cycle of $H_v - T_v$ of length at least 7 is an induced path of length at least 3 between two neighbors of v in $H_v - T_v$, and the shortest path between those vertices and the induced path have the same length, as $H_v - (T_v \cup \{v\})$ is distance-hereditary. Let x^* be an optimal fractional solution to Vertex Multicut, which can be efficiently found using the ellipsoid method and an algorithm for the (weighted) shortest path problem as a separation oracle. If $|x^*| \leq k$, then we can construct a multicut $X_v \subseteq V(H_v) \setminus (T_v \cup \{v\})$ of size $O(|x^*|^2) = O(k^2)$ using the approximation algorithm of Gupta [14]. If $|x^*| > k$, then any integral solution for $(H_v - (T_v \cup \{v\}), \mathcal{T})$ is larger than k, and any DH-modulator of size at most k must contain v. In this case, we add v to U.

We can confirm that $(G - U, k - |U|)$ is an instance equivalent to (G, k) and $S \cup (\bigcup_{v \in S \setminus U} (T_v \cup X_v))$ is a good DH-modulator for $G - U$. $\square$

5 Twin Reduction Rule

In a distance-hereditary graph, there may be a large set of pairwise twins. We introduce a reduction rule that bounds the size of a set of pairwise twins in $G-S$ by $\mathcal{O}(k^2|S|^3)$, where S is a DH-modulator (not necessarily good). The underlying observation is that it suffices to keep up to $k+1$ vertices that are pairwise twins with respect to each subset of S of small size. For a subset $S' \subseteq S$, two vertices u and v in $V(G) \setminus S$ are *S'-twins* if u and v have the same neighbors in S'. It is not difficult to get an upper bound $\mathcal{O}(k|S|^5)$, by considering all subsets S' of S of size $\min\{|S|, 5\}$ and marking up to $k+1$ S'-twins. To get a better bound, we proceed as follows.

Reduction Rule 1 *Let W be a set of pairwise twins in $G - S$, and let $m := \min\{|S|, 3\}$. (1) Over all subsets $S' \subseteq S$ of size m, we mark up to $k+1$ pairwise S'-twins in W that are unmarked yet. (2) When $|S| \geq 4$, over all subsets $S' \subseteq S$ of size 4, if there is an unmarked vertex v of W such that $G[S' \cup \{v\}]$ is isomorphic to the house or the gem, then we mark up to $k+1$ previously unmarked vertices in W including v that are pairwise S'-twins. (3) If there is an unmarked vertex v of W after finishing the marking procedure, we remove v from G.*

If (G, k) is irreducible with respect to Reduction Rule 1 and $|S| \geq 4$, then each set W of pairwise twins in $G - S$ contains $\mathcal{O}(k^2|S|^3)$ vertices, or (G, k) is a No-instance. This is because, if (G, k) is a Yes-instance, then all chosen subsets S' in (2) can be covered by at most $4k$ vertices. If $|S| \leq 3$, then W has at most $8(k+1)$ vertices. To see the safeness, suppose there is an unmarked vertex v of W after finishing the marking procedure. It is clear that if (G, k) is a Yes-instance, then $(G - v, k)$ is a Yes-instance. Suppose $G - v$ has a DH-modulator T of size at most k, and $G - T$ contains a DH obstruction F containing v. In case when $F - v$ is an induced path, let w, z be the end vertices of the path, and choose a set $S' \subseteq S$ of size 3 containing $\{w, z\} \cap S$. Since v is unmarked in Reduction Rule 1, there are $v_1, \ldots, v_{k+1} \in W \setminus \{v\}$ where $v_1, \ldots, v_{k+1}, v$ are pairwise S'-twins. Note that $V(F) \cap \{v_1, \ldots, v_{k+1}\} = \emptyset$ since no other vertex in F is adjacent to both w and z. Thus, there exists $v' \in \{v_1, \ldots, v_{k+1}\} \setminus T$ such that $G[V(F) \setminus \{v\} \cup \{v'\}]$ is a DH obstruction in $(G - v) - T$, contradiction.

If $|S \cap V(F)| \leq 3$, then we can proceed in the same way. We may assume $F - v$ is not an induced path and $|S \cap V(F)| \geq 4$. If F is the house or the gem, then we marked necessary vertices in (2), and thus we can proceed similarly. If F is the domino, then v should be a vertex of degree 2 in F. We can prove that the 3 vertices S' in $F - v$, two neighbors of v and the vertex farthest from v, satisfies that the existence of S'-twins with v is enough to get another DH obstruction.

6 The number of non-trivial components of $G - S$

We provide a reduction rule that bounds the number of non-trivial components of $G - S$, when S is a good DH-modulator. For $v \in S$ and a component C of $G - S$, let $N(v, C) := N_G(v) \cap V(C)$. We say that a pair (v, w) of vertices in S is a *witnessing pair* (for being non-split) for a component C of $G - S$ if $N(v, C) \neq \emptyset$, $N(w, C) \neq \emptyset$ and $N(v, C) \neq N(w, C)$. The following lemma is essential.

Lemma 2. *If $C_1, C_2, \ldots, C_m$ are distinct connected components of $G - S$ with $m \geq 2$ and $v_1, v_2, \ldots, v_m$ are distinct vertices of S ($v_{m+1} = v_1$) such that for each $i \in \{1, \ldots, m\}$, (v_i, v_{i+1}) is a witnessing pair for C_i, then $G[\{v_1, v_2, \ldots, v_m\} \cup \bigcup_{i \in \{1,\ldots,m\}} V(C_i)]$ contains a DH obstruction.*

Lemma 2 for $m = 2$ observes that if a pair of vertices in S witnesses at least $k+2$ non-trivial components in $G - S$, at least one of the pair must be contained in any size-k DH-modulator. Furthermore, keeping exactly $k+2$ non-trivial components would suffice to impose this restriction. This suggests the following rule.

Reduction Rule 2 *For each pair of vertices v and w in S, we mark up to $k+2$ non-trivial (previously unmarked) connected components C of $G - S$ such that (v, w) is a witnessing pair for C. If there is an unmarked non-trivial connected component C after the marking procedure, then we remove all edges in C.*

For the safeness of Reduction Rule 2, suppose there was an unmarked non-trivial connected component C after the marking procedure, and G' is the resulting graph. We mainly observe that if G' has a DH-modulator T of size at most k, then $(V(C) \setminus T, V(G) \setminus V(C) \setminus T)$ is a split in $G - T$. Otherwise, there are $v, w \in S$ and components $C_1, \ldots, C_{k+2}$ where (v, w) is a witnessing pair for $C, C_1, \ldots, C_{k+2}$ in G. Then there are 2 components among $C_1, \ldots, C_{k+2}$ that does not intersect T, and by Lemma 2, $G - T$ contains a DH obstruction, contradiction. Thus, if there is a DH obstruction H in $G - T$, then since $(V(C) \setminus T, V(G) \setminus V(C) \setminus T)$ is a split in $G - T$ and S is a good-modulator, we have $|V(H) \cap V(C)| \leq 1$. This implies that $G' - T$ also contains H, contradiction. We prove for the other direction in the similar way.

Proposition 3. *If (G, k) is irreducible with respect to Reduction Rule 2, then either the number of non-trivial components is $O(k^2|S|)$ or it is a* No*-instance.*

Proof (of Proposition 3). Suppose (G, k) is a YES-instance. We define an auxiliary multigraph F on S such that for $v, w \in S$, the multiplicity of the edge vw equals the number of non-trivial components that are marked by the witness of (v, w) in Reduction Rule 2. It suffices to obtain a bound on the number of edges in F with the edge multiplicity taken into account.

Construct a maximal packing of 2-cycles in F and let $S_1 \subseteq S$ be the vertices contained in the packing. By Lemma 2, a packing of size $k + 1$ implies the existence of $k + 1$ vertex-disjoint DH obstructions. Therefore, $|S_1| \leq 2k$. Again, due to the assumption that (G, k) is a YES-instance, the subgraph $F - S_1$ does not have $k + 1$ vertex-disjoint cycles: otherwise, G contains $k + 1$ vertex-disjoint DH obstructions by Lemma 2. By the Erdős-Pósa property of cycles, there exists $S_2 \subseteq V(F) \setminus S_1$ hitting all cycles of $F - S_1$ with $|S_2| \leq rk \log k$ for some constant r. Now, the number of edges in F is at most $|S_1| \cdot |S|(k + 2) + |S_2| \cdot |S \setminus S_1| + (|S \setminus S_1 \setminus S_2|) = 2k(k + 2)|S| + rk \log k|S| + |S| \leq (7 + r)k^2|S|$. $\square$

7 The size of non-trivial components of $G - S$

It remains to bound the size of each non-trivial connected component of $G - S$. For this, we need to use split decompositions that present tree-like structure of distance-hereditary graphs. For the length constraint, we shortly define here with an example, and put the full description in the appendix (preliminary section).

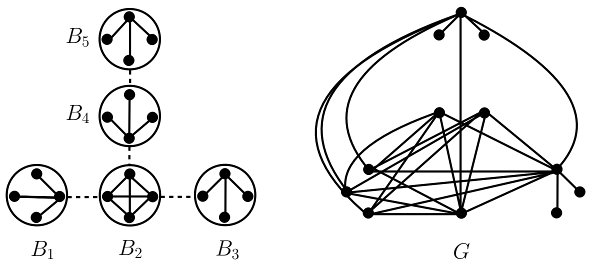

Fig. 2. An example of a split decomposition of a distance-hereditary graph. Dashed edges denote marked edges and each B_i denotes a bag.

A connected graph G is *prime* if $|V(G)| \geq 5$ and it has no split. A connected graph D with a distinguished set of cut edges $M(D)$ of D is called a *marked graph* if $M(D)$ forms a matching. An edge in $M(D)$ is a *marked edge*, and every other edge is an *unmarked edge*. A vertex incident with a marked edge is a *marked vertex*, and every other vertex is an *unmarked vertex*. Each component of $D - M(D)$ is a *bag* of D. See Figure 2 for an example. When G admits a split (A, B), we construct a marked graph D on the vertex set $V(G) \cup \{a', b'\}$ such that (1) $a'b'$ is a new marked edge, (2) there are no edges between A and B, (3) $\{a'\}$ is complete to $N_G(B)$, $\{b'\}$ is complete to $N_G(A)$, and (4) $G[A] = D[A]$ and $G[B] = D[B]$. The marked graph D is a *simple decomposition of* G. A *split decomposition* of a connected graph G is a marked graph D defined inductively to be either G or a marked graph defined from a split decomposition D' of G by replacing a bag B with its simple decomposition.

Cunningham and Edmonds [6] developed a canonical way to decompose a graph into a split decomposition. A split decomposition D of G is *canonical* if each bag of D is either a prime graph, a star, or a complete graph, and recomposing any marked edge of D violates this property. It is unique up to isomorphism [6] and can be computed in time $\mathcal{O}(|V(G)| + |E(G)|)$ [7]. In particular, Bouchet [3] proved that a graph is distance-hereditary if and only if every bag in its canonical split decomposition is a star or a complete graph.

Let D be the canonical split decomposition of a non-trivial component H of $G-S$. It is known that unmarked vertices in each bag of D consist of at most two twin sets in $G - S$. Thus, by Reduction Rule 1, it suffices to bound the number of bags of D. Since S is a good DH-modulator, for each $v \in S$, $G[V(H) \cup \{v\}]$ is distance-hereditary. Gioan and Paul [13, Theorem 3.4] described the way of extending D to the canonical split decomposition of $G[V(H) \cup \{v\}]$. In particular, there exists a bag or a marked edge that is modified when pushing v, and we can find this place in time $\mathcal{O}(|V(G)|)$. Such a bag or a marked edge is called *S-affected*. A bag B is a *branch bag* if $D - V(B)$ contains at least 3 connected components having at least two bags. For two adjacent bags B_1 and B_2, we denote by $e(B_1, B_2)$ the marked edge linking B_1 and B_2.

We first apply three reduction rules dealing with leaf bags. Firstly, we remove a vertex of degree 1 in G. Since any DH obstruction does not contain a vertex of degree 1, this is safe. Secondly, if there are a leaf bag B and its neighbor bag B' such that $B, B', e(B, B')$ are S-unaffected, and B' is a star bag whose center is adjacent to B, and $D - V(B')$ has exactly two components, then we remove all unmarked vertices of B'. In this case, all neighbors of a vertex in B' are contained in B whose unmarked vertices are pairwise twins in G. Thus, any DH obstruction does not contain a vertex of B', and this rule is safe. Lastly, if $A \subseteq V(H)$ is a maximal set of pairwise twins in G and flipping the adjacency between every two vertices of A reduces the number of bags, then we flip the adjacency. For instance, if B is a leaf bag that is a complete graph and its neighbor bag is a star bag whose leaf is adjacent to B, and B is S-unaffected, then by flipping the adjacency between two vertices in B, we can transform B into a star bag and merge with B'. This rule is also used in the FPT algorithm [9].

After applying those rules exhaustively, to find a reducible part, we color the bags of D with red and blue in the following way. (1) If a bag B is S-affected or incident with an S-affected edge, we color B with red. (2) If a bag B is adjacent to an S-affected leaf bag, we color B with red. (3) If B is a branch bag, then we color B with red. (4) All other bags are colored with blue. We can show that the number of red bags is at most $3|S|$, and for each bag B, there is at most one blue bag adjacent to B. We further show that the number of components of $D - \bigcup_{B \in \mathcal{R} \cup \mathcal{Q}} V(B)$ is at most $3|S|$, where $\mathcal{R}$ is the set of red bags, and $\mathcal{Q}$ is the set of blue leaf bags whose neighbor bags are red.

It remains to bound each connected component of $D - \bigcup_{B \in \mathcal{R} \cup \mathcal{Q}} V(B)$. Let D' be a connected component of $D - \bigcup_{B \in \mathcal{R} \cup \mathcal{Q}} V(B)$. As $\mathcal{R}$ contains all branch bags, D' contains no branch bags. Therefore, there is a sequence $B_1 - B_2 - \cdots - B_m$ of bags of D' that are not leaf bags, and all other bags are leaf bags adjacent to one of $B_1, \ldots, B_m$. For $1 \le i_1 < i_2 < i_3 \le m$, B_{i_2} is a (B_{i_1}, B_{i_3})-separator bag, if it is a star bag whose center is adjacent to neither B_{i_2-1} nor B_{i_2+1}. We first bound the number of (B_1, B_m)-separator bags, since if there are many, then we can merge two closest separator bags into one bag. This is similar to usual bypassing rule in Feedback Vertex Set. One interesting part is to bound the number of the sequence of consecutive bags that are not (B_1, B_m)-separator bags. We show that if there is such a sequence of length more than $5k+11$, then we can always find a vertex that can be safely removed. In the end, we reduce the number of bags in each component of $D - \bigcup_{B \in \mathcal{R} \cup \mathcal{Q}} V(B)$ to $20k + 52$, and we conclude that the number of all bags is bounded by $3|S|(20k + 54)$ bags.

Proof (of Theorem 1). We first prove that given an instance (G, k) and a good DH-modulator S, one can output an equivalent instance of size $O(k^5|S|^5)$. We first apply Reduction Rule 2 to (G, k) with S. After that, $G - S$ has $O(k^2|S|)$ non-trivial components or we can correctly report that (G, k) is a No-instance by Proposition 3. Next, we apply reduction rules for reducing the size of non-trivial components of $G-S$. We prove the safeness, polynomial-time applicability, and preserving the goodness of S in the appendix. In the end, the canonical split decomposition D of each non-trivial component of $G - S$ has at most $3|S|(20k+54)$ bags. Last, we apply Reduction Rule 1 exhaustively in polynomial time. This bounds the size of a twin set in $G-S$ by $O(k^2|S|^3)$. We note that the unmarked vertices of a bag form at most two twin sets. Therefore, the number of unmarked vertices in a bag is bounded by $O(k^2|S|^3)$. Especially, the same bounds apply to the number of trivial components in $G - S$ since they form an independent set in $G - S$. Combining the previous bounds altogether, we conclude that $V(G') = O(k^5|S|^5)$, when (G', k') is the resulting instance.

We may assume $n \le 2^{ck}$ for some constant c. Recall that there is an algorithm for DH Vertex Deletion running in time $2^{ck} n^{\mathcal{O}(1)}$ by Eiben, Ganian, and Kwon [9]. If $n > 2^{ck}$, then the algorithm of [9] solves the instance (G, k) correctly in polynomial time, in which case we can output a trivial equivalent instance. By Theorem 3, we can obtain a good DH-modulator S of size $O(k^5 \log n) = O(k^6)$ in polynomial time or correctly report (G, k) as a No-instance. The previous argument yields that in polynomial time, an equivalent instance (G', k') of size $O(k^{35})$ can be constructed. Now, applying Theorem 3 again[3] to (G', k'), we can either correctly conclude that (G', k'), and thus (G, k), is a No-instance or output a good DH-modulator S' of size $O(k^5 \log k)$. Finally we obtain a kernel of size $O(k^{30} \log^5 k)$.

References

1. Agrawal, A., Lokshtanov, D., Misra, P., Saurabh, S., Zehavi, M.: Feedback vertex set inspired kernel for chordal vertex deletion. In: Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19. pp. 1383–1398 (2017)

[3] That applying kernelization twice can yield an improved bound was adequately observed in [1].

2. Bandelt, H.J., Mulder, H.M.: Distance-hereditary graphs. J. Combin. Theory Ser. B 41(2), 182–208 (1986)
3. Bouchet, A.: Transforming trees by successive local complementations. J. Graph Theory 12(2), 195–207 (1988)
4. Courcelle, B., Makowsky, J.A., Rotics, U.: Linear time solvable optimization problems on graphs of bounded clique-width. Theory Comput. Syst. 33(2), 125–150 (2000)
5. Courcelle, B., Olariu, S.: Upper bounds to the clique width of graphs. Discrete Appl. Math. 101(1-3), 77–114 (2000)
6. Cunningham, W.H., Edmonds, J.: A combinatorial decomposition theory. Canad. J. Math. 32(3), 734–765 (1980)
7. Dahlhaus, E.: Parallel algorithms for hierarchical clustering, and applications to split decomposition and parity graph recognition. Journal of Algorithms 36(2), 205–240 (2000)
8. Diestel, R.: Graph Theory. No. 173 in Graduate Texts in Mathematics, Springer, third edn. (2005)
9. Eiben, E., Ganian, R., Kwon, O.: A Single-Exponential Fixed-Parameter Algorithm for Distance-Hereditary Vertex Deletion. In: Faliszewski, P., Muscholl, A., Niedermeier, R. (eds.) 41st International Symposium on Mathematical Foundations of Computer Science, MFCS 2016. Leibniz International Proceedings in Informatics (LIPIcs), vol. 58, pp. 34:1–34:14. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2016)
10. Feige, U., Hajiaghayi, M., Lee, J.R.: Improved approximation algorithms for minimum weight vertex separators. SIAM J. Comput. 38(2), 629–657 (2008)
11. Fomin, F.V., Lokshtanov, D., Misra, N., Saurabh, S.: Planar F-Deletion: approximation, kernelization and optimal FPT algorithms. In: 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science—FOCS 2012, pp. 470–479. IEEE Computer Soc., Los Alamitos, CA (2012)
12. Giannopoulou, A.C., Jansen, B.M.P., Lokshtanov, D., Saurabh, S.: Uniform kernelization complexity of hitting forbidden minors. In: Automata, languages, and programming. Part I, Lecture Notes in Comput. Sci., vol. 9134, pp. 629–641. Springer, Heidelberg (2015)
13. Gioan, E., Paul, C.: Split decomposition and graph-labelled trees: characterizations and fully dynamic algorithms for totally decomposable graphs. Discrete Appl. Math. 160(6), 708–733 (2012)
14. Gupta, A.: Improved results for directed multicut. In: Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2003, January 12-14, 2003, Baltimore, Maryland, USA. pp. 454–455 (2003)
15. Jansen, B.M.P., Pilipczuk, M.: Approximation and kernelization for chordal vertex deletion. In: 28th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, January 16-19, 2017, Barcelona, Spain. (2017), to appear
16. Kanté, M.M., Kim, E.J., Kwon, O., Paul, C.: An fpt algorithm and a polynomial kernel for linear rankwidth-1 vertex deletion. Algorithmica pp. 1–30 (2016)
17. Kim, E.J., Langer, A., Paul, C., Reidl, F., Rossmanith, P., Sau, I., Sikdar, S.: Linear kernels and single-exponential algorithms via protrusion decompositions. ACM Trans. Algorithms 12(2), Art. 21, 41 (2016)
18. Lewis, J.M., Yannakakis, M.: The node-deletion problem for hereditary properties is NP-complete. J. Comput. Syst. Sci. 20(2), 219–230 (1980)
19. Oum, S.: Rank-width and vertex-minors. J. Combin. Theory Ser. B 95(1), 79–100 (2005)