# Kernel Bounds for Structural Parameterizations of Pathwidth*

Hans L. Bodlaender, Bart M.P. Jansen, and Stefan Kratsch

Utrecht University, The Netherlands
{h.l.bodlaender,b.m.p.jansen,s.kratsch}@uu.nl

**Abstract.** Assuming the AND-distillation conjecture, the PATHWIDTH problem of determining whether a given graph $G$ has pathwidth at most $k$ admits no polynomial kernelization with respect to $k$. The present work studies the existence of polynomial kernels for PATHWIDTH with respect to other, structural, parameters.

Our main result is that, unless NP $\subseteq$ coNP/poly, PATHWIDTH admits no polynomial kernelization even when parameterized by the vertex deletion distance to a clique, by giving a cross-composition from CUTWIDTH. The cross-composition works also for TREEWIDTH, improving over previous lower bounds by the present authors. For PATHWIDTH, our result rules out polynomial kernels with respect to the distance to various classes of polynomial-time solvable inputs, like interval or cluster graphs.

This leads to the question whether there are nontrivial structural parameters for which PATHWIDTH does admit a polynomial kernelization. To answer this, we give a collection of graph reduction rules that are safe for PATHWIDTH. We analyze the success of these results and obtain polynomial kernelizations with respect to the following parameters: the size of a vertex cover of the graph, the vertex deletion distance to a graph where each connected component is a star, and the vertex deletion distance to a graph where each connected component has at most $c$ vertices.

## 1 Introduction

The notion of kernelization provides a systematic way to mathematically analyze what can be achieved by (polynomial-time) preprocessing of combinatorial problems [1]. This paper discusses kernelization for the problem to determine the *pathwidth* of a graph. The notion of pathwidth was introduced by Robertson and Seymour in their fundamental work on graph minors [2], and is strongly related to the notion of treewidth. There are several notions that are equivalent to pathwidth including *interval thickness*, *vertex separation number*, and *node search number* (see [3] for an overview). The problem to determine the pathwidth of a graph is well studied, also under the different names of the problem.

* This work was supported by the Netherlands Organization for Scientific Research (N.W.O.), project "KERNELS: Combinatorial Analysis of Data Reduction".

F.V. Fomin and P. Kaski (Eds.): SWAT 2012, LNCS 7357, pp. 352–363, 2012.

It is well known that the decision problem corresponding to pathwidth is NP-complete, even on restricted graph classes such as bipartite graphs and chordal graphs [4,5]. A commonly employed practical technique is therefore to preprocess the input before trying to compute the pathwidth, by employing a set of (reversible) data reduction rules. Similar preprocessing techniques for the TREEWIDTH problem have been studied in detail [6,7], and their practical use has been verified in experiments [8]. Using the concept of kernelization we may analyze the quality of such preprocessing procedures within the framework of parameterized complexity. A *parameterized problem* is a language $Q \subseteq \Sigma^* \times \mathbb{N}$, and such a problem is (strongly uniform) *fixed-parameter tractable* (FPT) if there is an algorithm that decides membership of an instance $(x, k)$ in time $f(k)|x|^{\mathcal{O}(1)}$ for some computable function $f$. A *kernelization* (or *kernel*) for $Q$ is a polynomial-time algorithm which transforms each input $(x, k)$ into an *equivalent* instance $(x', k')$ such that $|x'|, k' \leq g(k)$ for some computable function $g$, which is the *size* of the kernel. Kernels of polynomial size are of particular interest due to their practical applications. To analyze the quality of preprocessing rules for PATHWIDTH we therefore study whether they yield polynomial kernels for suitable parameterizations of the PATHWIDTH problem.

As the pathwidth of a graph equals the maximum of the pathwidth of its connected components, the PATHWIDTH problem with standard parameterization is AND-compositional and thus has no polynomial kernel unless the AND-distillation conjecture does not hold [9]. We thus do not expect to have kernels for PATHWIDTH of size polynomial in the target value for pathwidth $k$, and we consider whether polynomial kernels can be obtained with respect to other parameterizations.

As PATHWIDTH is known to be polynomial-time solvable when restricted graph classes such as interval graphs [3], trees [10] and cographs [11], it seems reasonable to think that determining the pathwidth of a graph $G$ which is "almost" an interval graph should also be polynomial-time solvable. Formalizing the notion of "almost" as the number of vertices that have to be deleted to obtain a graph in the restricted class $\mathcal{F}$, we can study the extent to which data reduction is possible for graphs which are close to polynomial-time solvable instances through the following problem:

PATHWIDTH PARAMETERIZED BY A MODULATOR TO $\mathcal{F}$
**Instance:** A graph $G = (V, E)$, a positive integer $k$, and a set $S \subseteq V$ such that $G - S \in \mathcal{F}$.
**Parameter:** $\ell := |S|$.
**Question:** $\mathbf{pw}(G) \leq k$?

The set $S$ is a *modulator* to the class $\mathcal{F}$. Observe that pathwidth should be polynomial-time solvable on $\mathcal{F}$ in order for this parameterized problem to be FPT. Our main result is a kernel lower bound for such a parameterization of PATHWIDTH. We prove that despite the fact that the pathwidth of an interval graph is simply the size of its largest clique minus one — which is very easy to find on interval graphs — the PATHWIDTH problem parameterized by a modulator to an interval graph does not admit a polynomial kernel unless

NP $\subseteq$ coNP/poly. In fact, we prove the stronger statement that, under the same condition, Pathwidth parameterized by a modulator to a single clique (i.e., by distance to $\mathcal{F}$ consisting of all complete graphs) does not admit a polynomial kernel[1] (Section 5). As the graph resulting from the lower-bound construction is co-bipartite, its pathwidth and treewidth coincide [12]: a corollary to our theorem therefore shows that Treewidth parameterized by vertex-deletion distance to a clique does not admit a polynomial kernel unless NP $\subseteq$ coNP/poly, thereby significantly strengthening a result of our earlier work [6] where we only managed to prove kernel lower bounds by modulators from cluster graphs and co-cluster graphs.

Our kernel bound effectively shows that even in graphs which are cliques after the deletion of $k$ vertices, the information contained in the (non)edges between these $k$ vertices and the clique is such that we cannot decrease the size of the clique to polynomial in $k$ in polynomial time, without changing the answer in some cases.

Faced with these negative results, we try to formulate *safe* reduction rules for Pathwidth (Section 3). It turns out that many of the rules for Treewidth (e.g., the rules involving (almost) simplicial vertices) are invalid when applied to Pathwidth, and more careful reduction procedures are needed to reduce the number of such vertices. We obtain several reduction rules for pathwidth, and show that they lead to provable data reduction guarantees when analyzed using a suitable parameterization (Section 4). In particular we prove that Pathwidth parameterized by a vertex cover $S$ (i.e., using $\mathcal{F}$ as the class of edgeless graphs in the template above) admits a kernel with $\mathcal{O}(|S|^3)$ vertices, that the parameterization by a modulator $S'$ to a disjoint union of stars has a kernel with $\mathcal{O}(|S'|^4)$ vertices, and finally that parameterizing by a set $S''$ whose deletion leaves a graph in which every connected component has at most $c$ vertices admits a kernel with $\mathcal{O}(c \cdot |S''|^3 + c^2 \cdot |S''|^2)$ vertices.

## 2 Preliminaries

In this work all graphs are finite, simple, and undirected. The open neighborhood of a vertex $v \in V$ in a graph $G$ is denoted by $N_G(v)$, and its closed neighborhood is $N_G[v]$. For sets of vertices $W \subseteq V$ we let $N_G[W] = \bigcup_{v \in W} N_G[v]$ and $N_G(W) = N_G[W] \setminus W$. If $S \subseteq V$ is a vertex set then $G - S$ denotes the graph obtained from $G$ by deleting all vertices of $S$ and their incident edges. For a single vertex $v$ we write $G - v$ instead of $G - \{v\}$. A vertex $v$ is *simplicial* in a graph $G$ if $N_G(v)$ is a clique. A vertex $v \in V$ is *almost simplicial* in a graph $G$ if $v$ has a neighbor $w$ such that $N_G(v) - \{w\}$ is a clique. In such a case, we call $w$ the *special neighbor* of $v$. For a set of vertices $W \subseteq V$, the subgraph of $G$

[1] For completeness we point out that Pathwidth parameterized by a modulator to a clique is FPT: try all orderings in which the vertices from $S$ can be introduced and forgotten in a decomposition, and do a polynomial-time computation for each ordering to find the best way to fit the clique $G - X$ into the decomposition.

induced by $W$ is denoted as $G[W]$. A *path decomposition* of a graph $G = (V, E)$ is a non-empty sequence $(X_1, \ldots, X_r)$ of subsets of $V$ called *bags*, such that:

- $\bigcup_{1 \leq i \leq r} X_i = V$,
- for all edges $\{v, w\} \in E$ there is a bag $X_i$ containing $v$ and $w$, and
- for all vertices $v \in V$, the bags containing $v$ are consecutive in the sequence.

The *width* of a path decomposition is $\max_{1 \leq i \leq r} |X_i| - 1$. The *pathwidth* $\mathbf{pw}(G)$ of $G$ is the minimum width of a path decomposition of $G$. Throughout the paper we will often make use of the fact that the pathwidth of a graph does not increase when taking a minor. We also use the following results.

**Lemma 1 (Cf. [11]).** *If graph $G$ contains a clique $W$ then any path- or tree decomposition for $G$ has a bag containing all vertices of $W$.*

**Lemma 2.** *All graphs $G$ admit a minimum-width path decomposition in which each simplicial vertex is contained in exactly one bag of the decomposition.*

*Proof.* Lemma 1 shows that for each simplicial vertex $v$, any path decomposition of $G$ has a bag containing the clique $N[v]$. As removal of $v$ from all other bags preserves the validity of the decomposition, we may do so independently for all simplicial vertices to obtain a decomposition of the desired form. □

## 3 Reduction Rules

In this section we give a collection of reduction rules. Formally, each rule takes as input an instance $(G, S, k)$ of Pathwidth parameterized by a modulator to $\mathcal{F}$, and outputs an instance $(G', S', k')$. With the exception of occasionally outright deciding **yes** or **no**, none of our reduction rules change the modulator $S$ or the value of $k$. In the interest of readability we shall therefore be less formal in our exposition, and make no mention of the values of $S'$ and $k'$ in the remainder. We say that a rule is *safe for pathwidth* (or in short: safe) if for each input $(G, S, k)$ and output $(G', S', k')$, the pathwidth of $G$ is at most $k$ if and only if the pathwidth of $G'$ is at most $k'$. Any subset of the rules gives a 'safe' preprocessing algorithm for pathwidth: apply the rules until no longer possible. We will argue later that this takes polynomial time for our rules, and give kernel bounds for some parameters of the graphs.

### 3.1 Vertices of Small Degree

We start off with a few simple rules for vertices of small degree. Note that, necessarily, these rules are slightly more restrictive than for the treewidth case; e.g., we cannot simply delete vertices of degree one since trees have treewidth one but unbounded pathwidth. The first rule is trivial.

**Rule 1.** *Delete any vertex of degree zero.*

**Rule 2.** *If two degree-one vertices share their neighbor then delete one of them.*

Correctness of Rule 2 follows from insights on the pathwidth of trees, pioneered by Ellis et al. [10]. A self-contained proof is provided in the full version.

The following rule handles certain vertices of degree two; a correctness proof is given in the full version.

**Rule 3.** *Let $v$, $w$ be two vertices of degree two, and suppose $x$ and $y$ are common neighbors to $v$ and $w$ with $x \in S$. Then remove $w$ and add the edge $\{x, y\}$.*

### 3.2 Common Neighbors and Disjoint Paths

Rule 4 in this section also appears in our work on kernelization for treewidth [6] and traces back to well-known facts about treewidth (e.g.,[13,14]). It is also safe in the context of pathwidth; the safeness proof is identical to when dealing with treewidth and is hence deferred to the full version.

**Lemma 3.** *Let $v$ and $w$ be nonadjacent vertices. Suppose there are at least $k+1$ internally vertex disjoint paths from $v$ to $w$ in $(V, E)$. Then the pathwidth of $G$ is at most $k$, if and only if the pathwidth of $G' = (V, E \cup \{\{v, w\}\})$ is at most $k$.*

A special case of Lemma 3, and the implied Rule 4, is when $v$ and $w$ have at least $k+1$ common neighbors. As we do not want to increase the size of a modulator, we only add edges between pairs of vertices with at least one endpoint in the modulator; thus $G - S$ remains unchanged.

**Rule 4 (Disjoint paths (with a modulator)).** *Let $v \in S$ be nonadjacent to $w \in V$, and suppose there are at least $k+1$ paths from $v$ to $w$ that only intersect at $v$ and $w$, where $k$ denotes the target pathwidth. Then add the edge $\{v, w\}$.*

### 3.3 Simplicial Vertices

In this section, we give a safe rule that helps to bound the number of simplicial vertices of degree at least two in a graph. Recall that we already have rules for vertices of degree one and zero, which are trivially simplicial.

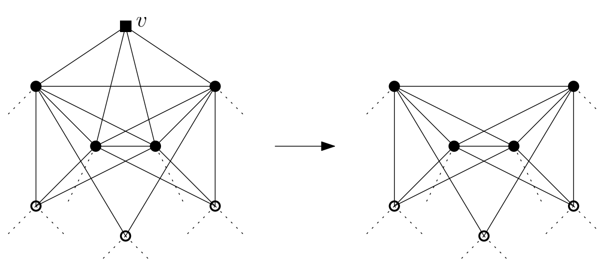


**Fig. 1.** An example of an application of the Simplicial vertex rule

**Lemma 4.** *Let $G = (V, E)$ be a graph, and let $v \in V$ be a simplicial vertex of degree at least two. If for all $x, y \in N_G(v)$ with $x \neq y$ there is a simplicial vertex $w \notin N_G[v]$ such that $x, y \in N_G(w)$, then* $\mathbf{pw}(G) = \mathbf{pw}(G - v)$*.*

*Proof.* As $G - v$ is a subgraph of $G$, we directly have that $\mathbf{pw}(G - v) \leq \mathbf{pw}(G)$. For the converse, let $(X_1, \ldots, X_r)$ be an optimal path decomposition of $G - v$. Using Lemma 2, we assume that for each simplicial vertex $x$, there is a unique bag $X_{i_x}$ with $N_G[x] \subseteq X_{i_x}$.

Let $C = N_G(v)$. A bag that contains $C$ is called a $C$-bag. As $C$ is a clique, Lemma 1 shows there is at least one $C$-bag. The $C$-bags must be consecutive in the path decomposition; let them be $X_{i_1}, \ldots, X_{i_2}$. We will first show there is a vertex $w \notin N_G[v]$ which is simplicial in $G - v$, and is contained in a $C$-bag. Let $x, y \in C$ (possibly with $x = y$) be vertices such that $x$ does not occur in bags with index smaller than $i_1$, and $y$ does not occur in bags of index larger than $i_2$.

If $x \neq y$ then let $w \notin N_G[v]$ be simplicial in $G$ such that $x, y \in N_G(w)$, whose existence is guaranteed by the preconditions. As $w$ is also simplicial in $G - v$ it occurs in a unique bag, which must be a $C$-bag since it must meet its neighbors $x$ and $y$ there. If $x = y$ then, as $v$ has degree at least two, there is a vertex $w \notin N_G[v]$ which is simplicial in $G$ and adjacent to $x$; hence its unique occurrence is also in a $C$-bag.

Thus we have established there is a vertex $w \notin N_G[v]$ which is simplicial in $G - v$ and is contained in exactly one bag, which is a $C$-bag $X_i$. Now insert a new bag just after $X_i$, with vertex set $X_i - \{w\} \cup \{v\}$. As $X_i - \{w\}$ contains all $v$'s neighbors, this gives a path decomposition of $G$ without increasing the width, and concludes the proof. □

Lemma 4 directly shows that Rule 5 is safe for Pathwidth.

**Rule 5.** *For each $e \in E$, compute span(e) as the number of simplicial vertices that are adjacent to both endpoints of the edge. If $v \in V$ is a simplicial vertex of degree at least two such that each edge between a pair of neighbors of $v$ has span at least 2, then remove $v$.*

### 3.4 Simplicial Components

Let $S$ be the set of vertices used as the modulator. We say that a set of vertices $W$ is a *simplicial component* if $W$ is a connected component in $G - S$ and $N_G(W) \cap S$ is a clique. Our next rule deals with simplicial components.

**Rule 6 (Simplicial components of known pathwidth).** *Let $S \subseteq V$ be the modulator and let $k$ denote the target pathwidth. Suppose that for each pair $v, w \in S \cap N_G(W)$ (including $v = w$), there are at least $2k+3$ simplicial components $Z \neq W$ such that $\{v, w\} \subseteq N_G(Z)$ and* $\mathbf{pw}(G[Z]) \geq \mathbf{pw}(G[W])$*. Then remove $W$ and its incident edges.*

Note that we have to include the case $v = w$ to ensure correctness for simplicial components which are adjacent to exactly one vertex in the modulator. The safeness proof for Rule 6 is given in the full version.

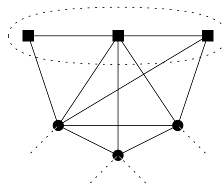

**Fig. 2.** The vertices marked with a square box form a simplicial component

Let us briefly discuss the running time of this reduction rule. As the modulator ensures that $G - S$ is contained in the graph class $\mathcal{F}$, the rule can be applied in polynomial time if the pathwidth of graphs in $\mathcal{F}$ can be determined efficiently. In the setting in which we apply the rule, the graphs in $\mathcal{F}$ are either disjoint unions of stars (which are restricted types of forests, allowing the use of the linear-time algorithm of Ellis et al. [10]), or $\mathcal{F}$ has constant pathwidth which means that the FPT algorithm for $k$-PATHWIDTH [13] runs in linear time.

### 3.5 Almost Simplicial Vertices

For almost simplicial vertices, we have a rule that replaces an almost simplicial vertex by a number of vertices of degree two. In several practical settings, the increase of number of vertices may be undesirable; the rule is useful to derive some theoretical bounds.

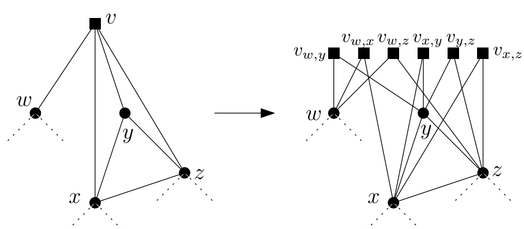


**Fig. 3.** An example of an application of the rule for almost simplicial vertices

**Lemma 5.** *Let $G = (V, E)$ be a graph and let $v \in V$ be an almost simplicial vertex of degree at least three, with special neighbor $w$. Let $G'$ be obtained by deleting $v$ and by adding a vertex $v_{p,q}$ with neighbors $p$ and $q$ for any $p, q \in N_G(v)$ with $p \neq q$. Then $\mathbf{pw}(G) = \mathbf{pw}(G')$.*

The proof of the lemma is postponed to the full version. The lemma justifies the following reduction rule, by observing that an almost simplicial vertex $v$ with $\deg_G(v) > k + 1$ means that $\mathbf{pw}(G) > k$, as $N_G[v] - w$ then forms a clique of size at least $k + 2$.

**Rule 7.** *Let $v \in V \setminus S$ be an almost simplicial vertex of degree at least three with special neighbor $w$. Let $k$ be the target pathwidth. If $\deg_G(v) > k+1$ then output **no**. Otherwise, delete $v$ and add a vertex $v_{p,q}$ with neighbors $p$ and $q$ for any $p, q \in N(v)$ with $p \neq q$.*

As a simplicial vertex is trivially almost simplicial, note that — in comparison to Rule 5 — the previous rule gives an alternative way of dealing with simplicial vertices.

## 4 Polynomial Kernelizations

For each of the safe rules given in the previous section, there is a polynomial time algorithm that tests if the rule can be applied, and if so, modifies the graph accordingly. (We assume that for Rule 6 the bound $\ell$ on the pathwidth of the components is a constant.) The following lemma shows that any algorithm that exhaustively applies (possibly just a subset of) these reduction rules can be implemented to run in polynomial time.

**Lemma 6.** *Each input instance $(G, S, k)$ is exhaustively reduced by $\mathcal{O}(n^2 + nk^2)$ applications of the reduction rules.*

*Proof.* First we note that for non-trivial instances, Rule 4 does not add edges to a vertex of degree at most two. In particular, no rule increases the number of vertices of degree at least three. So, we have at most $n$ applications of a rule that removes a vertex of degree at least three, and $\mathcal{O}(n^2)$ applications of Rule 4. Rule 7 is therefore executed at most $n$ times in total, and thus the number of vertices of degree two that are added in these steps is bounded by $\mathcal{O}(nk^2)$. As each other rule removes at least one vertex, the total number of rule applications in $G$ is bounded by $\mathcal{O}(n^2 + nk^2)$. □

By analyzing our reduction rules with respect to different structural parameters, we get the following results.

**Theorem 1.** Pathwidth parameterized by a modulator to $\mathcal{F}$ *admits polynomial kernels for the following choices of $\mathcal{F}$:*

1. *A kernel with $\mathcal{O}(\ell^3)$ vertices when $\mathcal{F}$ is the class of all independent sets, i.e., if the modulator $S$ is a vertex cover.*
2. *A kernel with $\mathcal{O}(c \cdot \ell^3 + c^2 \cdot \ell^2)$ vertices when $\mathcal{F}$ is the class of all graphs with connected components of size at most $c$.*
3. *A kernel with $\mathcal{O}(\ell^4)$ vertices when $\mathcal{F}$ is the class of all disjoint unions of stars.*

*Proof.* We show Part 3 followed by Part 2. Part 1 follows from the latter since it is a special case corresponding to $c = 1$.

**(Part 3.)** As stars have pathwidth one, graphs with a modulator $S$ of size $\ell$ to a set of stars have pathwidth at most $\ell + 1$. Thus, if $k \geq \ell + 1$, we return a dummy **yes**-instance of constant size. Now, assume $k \leq \ell$.

Our kernelization applies Rules 1–6 while possible, and applies Rule 7 to all vertices which have at most one neighbor in $G - S$. (Applying the rule to vertices with more neighbors in $G - S$ might cause the resulting graph $G' - S$ not to be a disjoint union of stars.) Recall for Rule 6 that $\mathbf{pw}(G - S) \leq 1$.

Let $(G, S, k)$ be a reduced instance. We will first bound the number of connected components of $G - S$, with separate arguments for simplicial and nonsimplicial components. Each component is a star, i.e., it is a single vertex or a $K_{1,r}$ for some $r$ (a center vertex with $r$ leaves). Note that in this proof the term leaf refers to a leaf of a star in $G - S$, independent of its degree in $G$ (and all degrees mentioned are with respect to $G$).

Associate each nonsimplicial component $C$ of $G - S$ to an arbitrary pair of nonadjacent neighbors of $C$ in $S$. It is easy to see that each such component provides a path between the two chosen neighbors, and that for different components these paths are internally vertex disjoint. Thus, since Rule 4 does not apply, no pair of vertices of $S$ has more than $k$ components associated to it. Hence there are at most $k \cdot |S|^2 = \mathcal{O}(\ell^3)$ nonsimplicial components.

Now consider a simplicial component $W$ of $G - S$, and note that $\mathbf{pw}(G[W]) \leq 1$. As Rule 6 does not apply, there is a pair $v, w \in S \cap N_G(W)$ (possibly $v = w$) such that there are strictly less than $2k + 3$ simplicial components $W' \neq W$ with $\mathbf{pw}(G[W']) \geq \mathbf{pw}(G[W])$ and $\{v, w\} \subseteq N_G(W')$. Associate $W$ to the pair $v, w$. It follows immediately that no pair of vertices of $S$ has more than $2k+3$ components associated to it, which gives a bound of $(2k + 3) \cdot |S|^2 = \mathcal{O}(\ell^3)$ on the number of simplicial components.

Thus we find that $G - S$ has a total of $\mathcal{O}(\ell^3)$ connected components (each of which is a star). This bounds the number of centers of stars by $\mathcal{O}(\ell^3)$. It remains to bound the total number of leaves that are adjacent to those centers.

Clearly, each star center has at most one leaf which has degree one (in $G$). Each leaf of degree two has exactly one neighbor in $S$ in addition to its adjacent star center. Since Rule 3 does not apply, no two leaves of degree two can have the same star center and neighbor in $S$; thus there are at most $\mathcal{O}(\ell^4)$ leaves of degree two.

Now, we are going to count the number of leaves (of stars) that are of degree more than two. For each such leaf, one neighbor is the center of its star and all other neighbors are in $S$. If its neighbors in $S$ would form a clique, then the leaf would be almost simplicial in $G$ (with the star center as the special neighbor) and Rule 7 would apply. Hence, as $G$ is reduced, we can associate each such leaf to a nonadjacent pair of vertices in $S$. As Rule 4 cannot be applied, we associate $\mathcal{O}(k)$ vertices to a pair, and thus the number of such leaves is bounded by $\mathcal{O}(k \cdot \ell^2) = \mathcal{O}(\ell^3)$.

Thus, the total number of vertices in $G$ is bounded by $\mathcal{O}(\ell^4)$. By Lemma 6 the reduction rules can exhaustively be applied in polynomial time. As the rules preserve the fact that $G - S$ is a disjoint union of stars, the resulting instance is a correct output for a kernelization algorithm. This completes the proof of Part 3.

**(Part 2.)** Fix some constant $c$ and let $\mathcal{F}$ be the class of all graphs of component size at most $c$. Let $(G, S, k)$ be an input instance. Note that the pathwidth of $G$ is bounded by $c+|S|-1$, since each component of $G-S$ has pathwidth at most $c-1$. We assume that $k \leq c+|S|-2$; otherwise the instance is **yes** and we may return a dummy **yes**-instance of constant size.

Our algorithm uses Rules 1, 2, 4, 5, and 6. Consider a graph $G$ where none of these rules can be applied. The bounds for the number of simplicial and nonsimplicial components of $G-S$ work analogously to Part 3; there are $\mathcal{O}(k|S|^2)$ components of the respective types. This gives a total of $\mathcal{O}(|S|+c\cdot(c+|S|)\cdot|S|^2) = \mathcal{O}(c^2|S|^2 + c|S|^3)$ vertices in $G$, using that $k \leq c+|S|-2$. This completes the proof of Part 2. □

## 5 Lower Bounds: Modulator to a Single Clique

We show that the problems Treewidth parameterized by a modulator to a single clique (TWMSC) and Pathwidth parameterized by a modulator to a single clique (PWMSC) do not admit a polynomial kernel unless NP ⊆ coNP/poly. In fact, we show that the results hold when restricted to co-bipartite graphs; as for these graphs the pathwidth equals the treewidth [12], the same proof works for both problems. The problems are covered by the general template given in the introduction, when using $\mathcal{F}$ as the class of all cliques.

To prove the lower bound we employ the technique of cross-composition [15], starting from the following NP-complete version [16, Corollary 2.10] of the Cutwidth problem:

Cutwidth on cubic graphs (CUTWIDTH3)
**Instance:** A graph $G$ on $n$ vertices in which each vertex has degree at least one and at most three, and an integer $k \leq |E(G)|$.
**Question:** Is there a linear layout of $G$ of cutwidth at most $k$, i.e., a permutation $\pi$ of $V(G)$ such that $\max_{i=1}^{n} |\{\{u,v\} \in E(G) \mid \pi(u) \leq i < \pi(v)\}| \leq k$?

As space restrictions prohibit us from presenting the full proof in this extended abstract, we will sketch the main ideas. To obtain a kernel lower bound through cross-composition, we have to embed the logical OR of a series of $t$ input instances of CUTWIDTH3 on $n$ vertices each into a single instance of the target problem for a parameter value polynomial in $n + \log t$. At the heart of our construction lies an idea of Arnborg et al. [4] employed in their NP-completeness proof for Treewidth. They interpreted the treewidth of a graph as the minimum cost of an elimination ordering on its vertices[2], and showed how for a given graph $G$ a co-bipartite graph $G^*$ can be created such that the cost of elimination orderings on $G^*$ corresponds to the cutwidth of $G$ under a related ordering.

[2] To eliminate a vertex in a graph means to remove it while completing its open neighborhood into a clique. When eliminating the vertices of a graph in the order given by $\pi$, the cost of the elimination ordering $\pi$ is the maximum degree of a vertex at the time it is eliminated.

We extend their construction significantly. By the degree bound, instances with $n$ vertices have $\mathcal{O}(n^2)$ different degree sequences. The framework of cross-composition thus allows us to work on instances with the same degree sequence (and same $k$). By enforcing that the structure of one side of the co-bipartite graph $G^*$ only has to depend on this sequence, all inputs can share the same "right hand side" of the co-bipartite graph; this part will remain small and act as the modulator. By a careful balancing act of weight values we then ensure that the cost of elimination orderings on the constructed graph $G^*$ are dominated by eliminating the vertices corresponding to exactly one of the input instances, ensuring that a sufficiently low treewidth is already achieved when one of the input instances is **yes**. On the other hand, the use of a binary-encoding representation of instance numbers ensures that low-cost elimination orderings for $G^*$ do not mix vertices corresponding to different input instances. The remaining details can be found in the full version of this paper. Our construction yields the following results.

**Theorem 2.** *Unless NP ⊆ coNP/poly,* PATHWIDTH *and* TREEWIDTH *do not admit polynomial kernels when parameterized by a modulator to a single clique.*

Interestingly, the parameter at hand is nothing else than the size of a vertex cover in the complement graph.

## 6 Conclusions

In this paper, we investigated the existence of polynomial kernelizations for PATHWIDTH. Taking into account that the problem is already known to be AND-compositional with respect to the target pathwidth — thus excluding polynomial kernels under the AND-distillation conjecture — we study alternative, structural parameterizations.

Our main result is that PATHWIDTH admits no polynomial kernelization with respect to the number of vertex deletions necessary to obtain a clique, unless NP ⊆ coNP/poly. This rules out polynomial kernels for vertex deletion distance from various interesting graph classes on which PATHWIDTH is known to be polynomial-time solvable, like chordal and interval graphs.

On the positive side we develop a collection of safe reduction rules for PATHWIDTH. Analyzing the effect of the rules we show that they give polynomial kernels with respect to the following parameters: vertex cover (i.e., distance from the class of independent sets), distance from graphs of bounded component size, and distance from disjoint union of stars.

It is an interesting open problem to determine whether there is a polynomial kernel for PATHWIDTH parameterized by the size of a feedback vertex set. For the related TREEWIDTH problem, a kernel with $\mathcal{O}(|S|^4)$ vertices is known [6], where $S$ denotes a feedback vertex set. Regarding PATHWIDTH, long paths in $G - S$ are the main obstacle that needs to be addressed by additional reduction rules.

## References

1. Guo, J., Niedermeier, R.: Invitation to data reduction and problem kernelization. SIGACT News 38(1), 31–45 (2007)
2. Robertson, N., Seymour, P.D.: Graph minors. I. Excluding a forest. J. Comb. Theory, Ser. B 35, 39–61 (1983)
3. Bodlaender, H.L.: A partial $k$-arboretum of graphs with bounded treewidth. Theor. Comput. Sci. 209(1-2), 1–45 (1998)
4. Arnborg, S., Corneil, D.G., Proskurowski, A.: Complexity of finding embeddings in a $k$-tree. SIAM Journal on Algebraic and Discrete Methods 8, 277–284 (1987)
5. Gustedt, J.: On the pathwidth of chordal graphs. Discrete Appl. Math. 45, 233–248 (1993)
6. Bodlaender, H.L., Jansen, B.M.P., Kratsch, S.: Preprocessing for Treewidth: A Combinatorial Analysis through Kernelization. In: Aceto, L., Henzinger, M., Sgall, J. (eds.) ICALP 2011, Part I. LNCS, vol. 6755, pp. 437–448. Springer, Heidelberg (2011)
7. van den Eijkhof, F., Bodlaender, H.L., Koster, A.M.C.A.: Safe reduction rules for weighted treewidth. Algorithmica 47(2), 139–158 (2007)
8. Bodlaender, H.L., Koster, A.M.C.A., van den Eijkhof, F.: Preprocessing rules for triangulation of probabilistic networks. Computational Intelligence 21(3), 286–305 (2005)
9. Bodlaender, H.L., Downey, R.G., Fellows, M.R., Hermelin, D.: On problems without polynomial kernels. J. Comput. Syst. Sci. 75(8), 423–434 (2009)
10. Ellis, J.A., Sudborough, I.H., Turner, J.S.: The vertex separation and search number of a graph. Inf. Comput. 113(1), 50–79 (1994)
11. Bodlaender, H.L., Möhring, R.H.: The pathwidth and treewidth of cographs. SIAM J. Discrete Math. 6, 181–188 (1993)
12. Möhring, R.H.: Triangulating graphs without asteroidal triples. Discrete Applied Mathematics 64(3), 281–287 (1996)
13. Bodlaender, H.L.: A linear-time algorithm for finding tree-decompositions of small treewidth. SIAM J. Comput. 25(6), 1305–1317 (1996)
14. Clautiaux, F., Carlier, J., Moukrim, A., Négre, S.: New Lower and Upper Bounds for Graph Treewidth. In: Jansen, K., Margraf, M., Mastrolli, M., Rolim, J.D.P. (eds.) WEA 2003. LNCS, vol. 2647, pp. 70–80. Springer, Heidelberg (2003)
15. Bodlaender, H.L., Jansen, B.M.P., Kratsch, S.: Cross-composition: A new technique for kernelization lower bounds. In: Proc. 28th STACS, pp. 165–176 (2011)
16. Monien, B., Sudborough, I.H.: Min cut is NP-complete for edge weighted trees. Theor. Comput. Sci. 58, 209–229 (1988)