A Cubic Vertex-Kernel for Trivially Perfect Editing

Maël Dumas

Univ. Orléans, INSA Centre Val de Loire, LIFO EA 4022, F-45067 Orléans, France

Anthony Perez

Univ. Orléans, INSA Centre Val de Loire, LIFO EA 4022, F-45067 Orléans, France

Ioan Todinca

Univ. Orléans, INSA Centre Val de Loire, LIFO EA 4022, F-45067 Orléans, France

— Abstract

We consider the Trivially Perfect Editing problem, where one is given an undirected graph G=(V,E) and a parameter $k\in\mathbb{N}$ and seeks to edit (add or delete) at most k edges from G to obtain a trivially perfect graph. The related Trivially Perfect Completion and Trivially Perfect Deletion problems are obtained by only allowing edge additions or edge deletions, respectively. Trivially perfect graphs are both chordal and cographs, and have applications related to the tree-depth width parameter and to social network analysis. All variants of the problem are known to be NP-complete [6, 29] and to admit so-called polynomial kernels [13, 23]. More precisely, the existence of an $O(k^3)$ vertex-kernel for Trivially Perfect Completion was announced by Guo [23] but without a stand-alone proof. More recently, Drange and Pilipczuk [13] provided $O(k^7)$ vertex-kernels for these problems and left open the existence of cubic vertex-kernels. In this work, we answer positively to this question for all three variants of the problem.

2012 ACM Subject Classification Theory of computation \rightarrow Parameterized complexity and exact algorithms

Keywords and phrases Parameterized complexity, kernelization algorithms, graph modification, trivially perfect graphs

Digital Object Identifier 10.4230/LIPIcs.MFCS.2021.45

Related Version Full Version: https://arxiv.org/abs/2105.08549

Introduction

A broad range of optimization problems on graphs are particular cases of so-called modification problems. Given an arbitrary graph G = (V, E) and an integer k, the question is whether G can be turned into a graph satisfying some desired property by at most k modifications. By modifications we mean, according to the problem, vertex deletions (as for Vertex Cover and Feedback Vertex Set where we aim to obtain graphs with no edges, or without cycles respectively) or edge deletions and/or additions (as for Minimum Fill-In, also known as Chordal Completion, where the goal is to obtain a chordal graph, with no induced cycles with four or more vertices, by adding at most k edges).

Here we consider edge modifications problems, that can be split in three categories, depending whether we allow only edge additions, only edge deletions, or both operations, in which case we speak of edge editing. Consider a family \mathcal{H} of graphs, called *obstructions*. In the \mathcal{H} -free edges of edge editing problem we seek to edit at most k edges of G to obtain a graph that does not contain any obstruction from \mathcal{H} as an induced subgraph. One can similarly define \mathcal{H} -free completion and \mathcal{H} -free deletion variants of this problem by only allowing the addition or deletion of edges, respectively. E.g., MINIMUM FILL-IN corresponds to \mathcal{H} -free completion, where \mathcal{H} is formed by all cycles with at least four vertices. For most families \mathcal{H} , all three versions are NP-complete, but thinking of k as of some suitably small quantity, they have been intensively studied in the framework of parameterized complexity (see [11] for

a comprehensive survey). The aim of parameterized complexity is to determine whether it is possible to decide the instance at hand in time $f(k) \cdot n^{O(1)}$ for some computable function f. Such problems are said to be FPT (fixed-parameter tractable). With a simple but elegant and powerful argument, Cai [7] proved that whenever \mathcal{H} is finite all three variants are FPT. Basically, whenever the graph contains one of the obstructions (graphs of \mathcal{H}), the algorithm branches on all possible modifications to destroy it, and makes the recursive calls with a lesser parameter k. When the family \mathcal{H} contains all cycles with at least four vertices, the corresponding edition problem Chordal Editing was shown to be FPT relatively recently [10]. The completion variant, i.e., the MINIMUM FILL-IN, was known to be FPT since the 90's [7, 25].

We consider an equivalent definition of fixed-parameter tractability, namely kernelization. Given a parameterized problem Π , a kernelization algorithm for Π (or kernel for short) is an algorithm that given any instance (I,k) of Π runs in time polynomial in |I| and k and outputs an equivalent instance (I', k') of Π such that $|I'| \leq h(k)$ and $k' \leq g(k)$ for some computable functions q and h. Whenever h is polynomial, we say that Π admits a polynomial kernel. A kernelization algorithm uses a set of polynomial-time computable reduction rules to reduce the instance at hand. We say that a reduction rule is safe whenever its application on an instance (I, k) of Π results in an equivalent instance (I', k') of Π . It is well-known that a parameterized problem is FPT if and only if it admits a kernelization algorithm [17]. While many polynomial kernels are known to exist for editing problems (see [11] or [27] for surveys), it is known that some editing problems are unlikely to admit polynomial kernels under reasonable complexity-theoretic assumptions [8, 22, 26]. When \mathcal{H} contains only a single obstruction, several results towards a dichotomy regarding the existence of polynomial kernels have been obtained [1, 8, 28]. Very recently, Marx and Sandeep [28] narrowed down the problem for obstructions containing at least 5 vertices to only nine distinct obstructions. In other words, the non-existence of polynomial kernels for \mathcal{H} -FREE EDITING for all such obstructions would imply the non-existence of polynomial kernels for any obstruction with at least 5 vertices. When \mathcal{H} contains several obstructions, a very natural setting is to include all cycles of length at least 3 in \mathcal{H} , thus targeting a subclass of chordal graphs. Indeed, editing (and especially completion) problems towards such classes cover classical problems with both theoretical and practical interest [15, 21, 24, 25, 33]. Notice that many known polynomial kernels for editing problems concern such classes [3, 4, 13, 23, 25]. For completion and deletion versions, polynomial kernels are often used as a first step in the design of subexponential parameterized algorithms [5, 12, 18, 19].

In this work, we focus on editing problems towards trivially perfect graphs, that is $\mathcal{H} = \{P_4, C_4\}$ (respectively a path and a cycle on 4 vertices). This problem is known as TRIVIALLY PERFECT EDITING in the literature. By allowing edge addition or edge deletion only, we obtain the TRIVIALLY PERFECT COMPLETION and TRIVIALLY PERFECT DELETION problems, respectively.

Related work

While the NP-Completeness of TRIVIALLY PERFECT COMPLETION and TRIVIALLY PERFECT DELETION has been known for some time [6], the complexity of TRIVIALLY PERFECT EDITING remained open until a work of Nastos and Gao [29]. Trivially perfect graphs have recently regained attention since they are related to the well-studied width parameter tree-depth [20, 30] which corresponds to the size of the largest clique of a trivially perfect supergraph of G with the smallest clique number. Moreover, Nastos and Gao [29] proposed a new definition for

community structure based on small obstructions. In particular, the authors emphasized that editing a given graph into a trivially perfect graph yields meaningful clusterings in real networks [29]. Trivially perfect graphs also correspond to chordal cographs and admit a so-called universal clique decomposition [12]. Polynomial kernels with $O(k^7)$ vertices have been obtained for all variants of the problem by Drange and Pilipczuk [13]. The technique used relies on a reduction rule bounding the number of vertices in any trivially perfect module and the computation of a so-called vertex modulator, that is a maximal packing of obstructions with additional properties. Combined with sunflower-like reduction rules and a careful analysis of the graph remaining apart from the vertex modulator, the authors managed to provide polynomial kernels. They then asked whether the $O(k^7)$ bound could be improved, and qualify as "really challenging question" whether one can match the $O(k^3)$ bound for Trivially Perfect Completion claimed by Guo [23].

Our contribution

We answer positively to this question and provide kernels with $O(k^3)$ vertices for all considered problems. To be complete, a quadratic vertex-kernel for the completion version only is claimed in [2, 9]. While our kernelization algorithm shares similarities with the work of Drange and Pilipczuk [13], our technique differs in several points. In particular, we do not rely on the computation of a vertex modulator, a useful technique to design polynomial kernels but somehow responsible for the large bound obtained. To circumvent this issue, we only rely on the so-called universal clique decomposition of trivially perfect graphs. This decomposition partitions the vertices of trivially perfect graph G into cliques, the bags being structured as nodes of a rooted forest such that two vertices are adjacent in G if and only they are in a same bag, or in two bags such that one is an ancestor of the other in the forest. For any positive instance of the problem, at most 2k bags contain vertices incident to modified edges. Informally, the rest of the bags can be regrouped into two types of 'chunks'. Some correspond to trivially perfect modules of the input graph (which are known to be reducible to small sizes by [13]), others have a more complicated but still particular structure, similar to the *combs* of [13]. We show how to reduce the size of these combs. Altogether we believe that our rules not only improve the size of the kernel but also significantly simplify the kernelization algorithm of [13]. Last but not least, we think that this approach based on tree-like decompositions and the analysis of large chunks of the graph that are not affected by the modified edges might be exploitable for other editing problems. Indeed the technique has strong similarities with the notion of branches introduced by Bessy et al. [3] for modification to 3-leaf power graphs, a closely related graph class.

Outline

We begin with some preliminaries definitions and results about trivially perfect graphs (Section 1). We then introduce the notion of combs and provide the set of reduction rules needed to obtain an $O(k^3)$ vertex-kernel for Trivially Perfect Editing (Section 2). The combinatorial bound on the kernel size is provided in Section 3. We explain how these results can be adapted to obtain similar kernels for Trivially Perfect Completion and Trivially Perfect Deletion (Section 4). The Conclusion section summarizes the results and suggests further developments. Proofs of statements labeled with (\star) are omitted in this extended abstract. The interested reader may refer to [14] for a full version of this paper.

1 Preliminaries

We consider simple, undirected graphs G = (V, E) where V denotes the vertex set and $E \subseteq (V \times V)$ the edge set of G. We will sometimes use V(G) and E(G) to clarify the context. Given a vertex $u \in V$, the open neighborhood of u is the set $N_G(u) = \{v \in V : uv \in E\}$. The closed neighborhood of u is defined as $N_G[u] = N_G(u) \cup \{u\}$. A vertex $u \in V$ is universal if $N_G[u] = V$, and two vertices u and v are true twins if $N_G[u] = N_G[v]$. The set of universal vertices forms a clique and is called the universal clique of G. Given a subset of vertices $S \subseteq V$, $N_G[S]$ is the set $\bigcup_{v \in S} N_G[v]$ and $N_G(S)$ is the set $N_G[S] \setminus S$. We will omit the mention to G whenever the context is clear. The subgraph induced by S is defined as $G[S] = (S, E_S)$ where $E_S = \{uv \in E : u \in S, v \in S\}$. For the sake of readability, given a subset $S \subseteq V$ we define $G \setminus S$ as $G[V \setminus S]$. A subset of vertices $C \subseteq V$ is a connected component of G if G[C] is a maximal connected subgraph of G. A subset of vertices $M \subseteq V$ is a module of G if and only if $N_G(u) \setminus M = N_G(v) \setminus M$ holds for every $u, v \in M$. A maximal set of true twins $K \subseteq V$ is a critical clique. Notice that G[K] is a clique module and that the set K(G) of critical cliques of any graph G partitions its vertex set V(G). Notice that the universal clique is a critical clique.

Trivially perfect graphs

A graph G = (V, E) is trivially perfect if and only if it does not contain any P_4 (a path on 4 vertices) nor C_4 (a cycle on 4 vertices) as an induced subgraph (see Figure 1). We consider the following problem.

TRIVIALLY PERFECT EDITING

Input: A graph G = (V, E), a parameter $k \in \mathbb{N}$

Question: Does there exist a set of pairs $F \subseteq (V \times V)$ of size at most k such that the graph $H = (V, E \triangle F)$ is trivially perfect, with $E \triangle F = (E \setminus F) \cup (F \setminus E)$?

Given an instance (G = (V, E), k) of Trivially Perfect Editing, a set $F \subseteq (V \times V)$ such that $H = (V, E \triangle F)$ is trivially perfect is an *edition* of G. When F is constrained to be disjoint from (resp. contained in) E, we say that F is a *completion* (resp. a *deletion*) of G. The corresponding problems are Trivially Perfect Completion and Trivially Perfect Deletion, respectively. For the sake of simplicity, given an edition (resp. completion, deletion) F of G, we use $G \triangle F$, G + F and G - F to denote the graphs $(V, E \triangle F)$, $(V, E \cup F)$ and $(V, E \setminus F)$, respectively. A vertex is affected by F whenever it is contained in some pair of F. The set F is a k-edition (resp. k-completion, k-deletion) whenever $|F| \leq k$. Finally, we say that such a set F is optimal whenever it is minimum-sized.

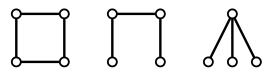


Figure 1 The C_4 , P_4 and claw graphs, respectively. The claw will be useful in some of our proofs.

Trivially perfect graphs are hereditary and closed under true twin addition. This property will be useful to deal with critical cliques, as stated by the following result. Recall that critical cliques are maximal sets of true twins (or, equivalently, maximal clique modules), they will play a central role throughout this paper.

▶ Lemma 1 ([3]). Let \mathcal{G} be a hereditary class of graphs closed under true twin addition. For every graph G = (V, E), there exists an optimal edition (resp. completion, deletion) F into a graph of \mathcal{G} such that for any two critical cliques K and K either $(K \times K') \subseteq F$ or $(K \times K') \cap F = \emptyset$.

Several characterizations are known to exist for trivially perfect graphs. We will mainly use the following ones.

- ▶ **Proposition 2** ([32]). The class of trivially perfect graphs can be defined recursively as follows:
- a single vertex is a trivially perfect graph.
- Adding a universal vertex to a trivially perfect graph results in a trivially perfect graph.
- The disjoint union of two trivially perfect graphs results in a trivially perfect graph.
- ▶ **Definition 3** (Universal clique decomposition, [12]). A universal clique decomposition (UCD) of a connected graph G = (V, E) is a pair $\mathcal{T} = (T = (V_T, E_T), \mathcal{B} = \{B_t\}_{t \in V_T})$ where T is a rooted tree and \mathcal{B} is a partition of the vertex set V into disjoint nonempty subsets, such that:
- if $vw \in E$ and $v \in B_t$, $w \in B_s$ then s and t are on a path from a leaf to the root, with possibly s = t, and
- for every node $t \in V_T$, the set of vertices B_t is the universal clique of the induced subgraph $G[\bigcup_{s \in V(T_t)} B_s]$, where T_t denotes the subtree of T rooted at t.

The vertices of T are called *nodes* of the decomposition, while the sets of \mathcal{B} are called *bags*. We will sometimes abuse notation and identify nodes of T with their corresponding bags in \mathcal{B} . Notice moreover that in a universal clique decomposition, every node t of T that is not a leaf has at least two children since otherwise B_t would not contain *all* universal vertices of $G[\bigcup_{s \in V(T_t)} B_s]$.

▶ Lemma 4 ([12]). A connected graph G admits a universal clique decomposition if and only if it is trivially perfect. Moreover, such a decomposition is unique up to isomorphisms.

One can observe that finding a universal clique decomposition can be done in polynomial time by iteratively identifying universal cliques and connected components. Finally, both Definition 3 and Lemma 4 can be naturally extended to disconnected trivially perfect graphs by considering a rooted forest instead of a rooted tree. More precisely, the universal clique decomposition of a disconnected graph G=(V,E) is a rooted forest of universal clique decompositions of its connected components. Such a graph is thus trivially perfect if and only if it admits a universal clique decomposition shaped like a rooted forest.

We conclude this section by providing a new characterization of trivially perfect graphs in terms of maximal cliques and nested families.

- ▶ **Definition 5** (Nested family). Let U be a universe and $\mathcal{F} \subseteq 2^{|U|}$ a family of subsets of U. The family \mathcal{F} is nested if and only if for every $A, B \in \mathcal{F}$, $A \subseteq B$ or $B \subseteq A$ holds.
- ▶ Lemma 6 (*). Let G = (V, E) be a graph, $S \subseteq V$ a maximal clique of G and $K_1, ..., K_r$ the connected components of $G \setminus S$. The graph G is trivially perfect if and only if the following conditions are verified:
 - (i) $G[S \cup K_i]$ is trivially perfect for every $1 \le i \le r$,
- (ii) $\bigcup_{1 \le i \le r} \{N_G(K_i)\}$ is a nested family,
- (iii) $(K_i \times N_G(K_i)) \subseteq E$ for every $1 \le i \le r$.

2 Kernelization algorithm for Trivially Perfect Editing

We begin this section by providing a high-level description of our kernelization algorithm. As mentioned in the introductory section, we use the universal clique decomposition of trivially perfect graphs to bound the number of vertices of a reduced instance. Let us consider a positive instance (G = (V, E), k) of Trivially Perfect Editing, F a suitable solution and $H = G \triangle F$. Denote by $\mathcal{T} = (T, \mathcal{B})$ the universal clique decomposition of H as described Definition 3. Since $|F| \leq k$, we know that at most 2k bags of \mathcal{T} may contain affected vertices. Let A be the set of such bags, and let A' denote the lowest common ancestor closure of A in forest T (Definition 17). As we shall see later, the size of A' is also linear in K (Lemma 18). The removal of every bag of A' from T will disconnect the forest T into several components (see Figure 2).

Such a connected component D of $T \setminus A'$ may see zero, one or two nodes of A' in the forest T (Lemma 4). If D has no neighbour in A', the union of all bags of D corresponds to a connected component of H and of G, inducing a trivially perfect graph in G, and will be eliminated by a reduction rule. We shall see that the union of all components D_a of the second type, seeing a unique bag $a \in A'$ in the forest T, corresponds to a trivially perfect module of graph G. We use the reductions rules of [13] to shrink such a module to $O(k^2)$ vertices, which boils down to a total $O(k^3)$ vertices since |A'| = O(k).

Our efforts will be focused on components D seeing two bags $a_1, a_2 \in A'$, one of them being ancestor of the other in forest T. We call such a structure D a comb (Definition 9 and Figure 2).

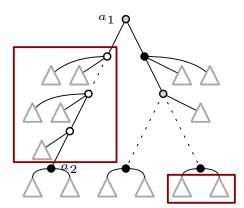


Figure 2 Analysis of a universal clique decomposition of a connected trivially perfect graph. Black vertices represent bags of A, gray vertices bags of A' and triangles are connected trivially perfect subgraphs of G. The leftmost rectangle is a comb of G, the rightmost a trivially perfect module. Note that any group of triangles rooted at a same bag is a trivially perfect module.

Such combs (the union of their bags) induce, in graph G, a trivially perfect subgraph that can be partitioned with regard to critical cliques and trivially perfect modules with nice inclusion properties on their neighborhoods. We provide two distinct reduction rules on these structures. Rule 4 reduces the so-called shaft of the comb (intuitively, the path strictly between a_1 and a_2 in T) to length O(k). Rule 5 reduces the size of the whole comb (the union of its bags) to $O(k^2)$. Altogether, the reduced instance cannot contain more than $O(k^3)$ vertices.

We would like to note that the combs considered in this work are similar to the ones defined by Drange and Pilipczuk [13] and thus named after them. However, the two structures are not strictly identical, in particular since they were originally defined with respect to a vertex modulator (i.e. a packing of obstructions), and thus their neighborhood towards the rest of the graph was structured differently.

In the remaining of this section we assume that we are given an instance (G = (V, E), k)of Trivially Perfect Editing.

2.1 Reducing critical cliques and trivially perfect modules

We first give a classical reduction rule when dealing with modification problems. This rule is trivially safe for trivially perfect graphs.

▶ Rule 1. Let $C \subseteq V$ be a subset of vertices such that G[C] is a trivially perfect connected component of G. Remove C from G.

We now give known reduction rules that deal with critical cliques and trivially perfect modules. The safeness of Rule 2 comes from the fact that trivially perfect graphs are hereditary and closed under true twin addition combined with Lemma 1. The safeness and polynomial-time application of Rule 3 was proved by Drange and Pilipczuk [13]. We would like to mention that while the statement of their rule assumes the instance at hand to be reduced by classical sunflower rules, this is actually not needed to prove the safeness of the rule. Altogether, we have the following.

- ▶ Rule 2. Let $K \subseteq V$ be a set of true twins of G such that |K| > k+1. Remove |K| (k+1)arbitrary vertices in K from G.
- ▶ Rule 3. Let $M \subseteq V$ be a module of G such that G[M] is trivially perfect and M contains an independent set I of size at least 2k + 5. Remove all vertices of $M \setminus I$ from G.
- ▶ Lemma 7 (Folklore, [3, 13]). Rules 1 to 3 are safe and can be applied in polynomial time.

Using a structural result on trivially perfect graphs where critical cliques and independent sets have bounded size, Drange and Pilipczuk [13] proved the following.

▶ Lemma 8 ([13]). Let (G = (V, E), k) be an instance of TRIVIALLY PERFECT EDITING reduced under Rules 2 and 3. Then for every module $M \subseteq V$ such that G[M] is trivially perfect, $|M| = O(k^2)$.

2.2 Reducing shafts of combs

We now consider the main structure of our kernelization algorithm, namely combs. Recall that such structures are similar to the ones defined by Drange and Pilipczuk [13] but not strictly identical. More precisely, the inner part of the structure is the same but not their neighborhoods towards the rest of the graph. We however choose to use the same name since it is well-suited to illustrate the structure (see Figure 3).

- ▶ **Definition 9** (Comb). Let G = (V, E) be a graph and $C, R \subseteq V$ be such that C is a clique which can be partitioned into l critical cliques $\{C_1,...,C_l\}$ and R can be partitioned into lnon-empty and non-adjacent trivially perfect modules $\{R_1,...,R_l\}$. The pair P=(C,R) is a comb if and only if:
- there exist $V_f, V_p \subseteq V(G) \setminus \{C, R\}, V_f \neq \emptyset$ such that $\forall x \in C, N_G(x) \setminus (C \cup R) = V_p \cup V_f$ and $\forall y \in R$, $N_G(y) \setminus (C \cup R) = V_p$, $N_G(C_i) \cap R = \bigcup_{j=i}^l R_j$ and $N_G(R_i) \cap C = \bigcup_{j=1}^i C_j$ for $1 \le i \le l$.

By the following property, given a comb (C, R) of graph G = (V, E), the subgraph $G[C \cup R]$ is trivially perfect, and has a universal clique decomposition in which critical cliques (C_1, \ldots, C_l) are arranged in a path starting from the root, the shaft of the comb, and the decomposition of



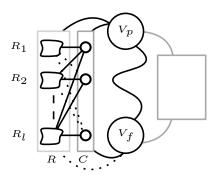


Figure 3 Illustration of a comb, with shaft C and teeth R. The edges between V_p and V_f can be anything. Every tooth R_i induces a (possibly disconnected) trivially perfect module.

each tooth R_i is attached to C_i ; see Figure 3. The length of (C,R) is l, the number of critical cliques in C. We can observe that $N_G[C_l] \subseteq \cdots \subseteq N_G[C_1]$ and $N_G(R_1) \subseteq \cdots \subseteq N_G(R_l)$ because for $1 \leq i \leq l$, $N_G[C_i] = (\bigcup_{j=i}^l R_j) \cup V_p \cup V_f$ and $N_G(R_i) = (\bigcup_{j=1}^i C_j) \cup V_p$.

- ▶ Proposition 10 (*). Given a comb (C,R) of graph G = (V,E), the subgraph $G[C \cup R]$ is trivially perfect. Moreover the sets V_p and V_f , and the ordered partitions (C_1, \ldots, C_l) of C and (R_1, \ldots, R_l) of R are uniquely determined.
- ▶ Lemma 11. Given an instance (G = (V, E), k) of TRIVIALLY PERFECT EDITING and a comb (C, R) of length $l \ge 2k + 2$ of G, there is no optimal k-edition that affects vertices in $C \cup R$.

Proof. Consider a k-edition F of G and $H = G \triangle F$. Denote by $F' \subseteq F$ the subset of pairs from F which does not contain any vertex from $C \cup R$ and let $H' = G \triangle F'$. Since $|F| \le k$ and (C, R) is a comb of length at least 2k + 2, there exist $i \ne j \in \{1, \ldots, l\}$ such that C_i, R_i, C_j and R_j do not include affected vertices of F. Let us take $c_1 \in C_i, r_1 \in R_i, c_2 \in C_j$ and $r_2 \in R_j$.

Suppose that H' is not trivially perfect, then there exists an obstruction W of H' such that $A = W \cap (C \cup R) \neq \emptyset$. Since pairs of F' do not contain vertices of $C \cup R$, (C, R)is a comb in H' and |A|=4 is impossible since $H'[C\cup R]=G[C\cup R]$ is trivially perfect by Proposition 10. We show that |A|=3 is also impossible. If |A|=3 then the vertex $x \in W \setminus (C \cup R)$ is in the set $V_p \cup V_f$, otherwise the obstruction W would not be connected. We now show that H'[W] is a claw, contains a triangle (as subgraph) or is not connected. If $x \in V_p$, then by construction x is adjacent to every vertex of the comb and H'[W] would be a claw. If $x \in V_f$ and A contains at least two vertices in C, then these vertices would induce a triangle with x. If $x \in V_f$ and A contains at least two vertices $r', r'' \in R$, then x is not adjacent to any of them (since V_f does not see R in G). If r' and r'' are not adjacent in H', either the fourth vertex of W sees r', r'' and x so H'[W] is a claw, or H'[W] is disconnected. If r' and r'' are adjacent in H', they must belong to a same module R_i . Again the fourth vertex of W must either see them both thus forming a triangle, or none of them and H'[W]is disconnected. In any case, A cannot be an obstruction and we conclude that either |A|=1or |A|=2. We shall now construct an obstruction $W'=(W\backslash A)\cup A'$ such that H'[W] and H'[W'] are isomorphic and $A' \subseteq \{c_1, r_1, c_2, r_2\}$. We can observe that W must contain a vertex from V_p or V_f .

If |A| = 1, take $x \in A$. If $x \in R$ then let $A' = \{r_1\}$, else let $A' = \{c_1\}$. Since (C, R) is a comb, H'[W] and H'[W'] are isomorphic.

If |A| = 2, denote by x and y the elements of A. If $x, y \in C$, then H[W] contains a triangle. If $x \in C$ and $y \in R$, in the subcase $xy \in E(H')$ let $A' = \{c_1, r_1\}$ and observe that $c_1r_1 \in E(H')$, hence H'[W] and H'[W'] are isomorphic; in the other subcase $xy \notin E(H')$, take $A' = \{c_2, r_1\}$, so $c_2r_1 \notin E(H')$ thus again H'[W] and H'[W'] are isomorphic. Eventually consider the last case $x, y \in R$. If $xy \in E(H')$ then H[W] contains a triangle, else $xy \notin E(H')$, so let $A' = \{r_2, r_1\}$ and note that $r_2r_1 \notin E(H')$ thus H'[W] and H'[W'] are isomorphic.

The set W' is an obstruction of H' and since the vertices in $\{c_1, r_1, c_2, r_2\}$ are not incident to any pair of F, W' is also an obstruction of H. Therefore H is not trivially perfect, which is a contradiction, concluding the proof of the Lemma.

- ▶ Rule 4. Given a comb (C, R) of length $l \ge 2k + 2$ of G, remove from G the vertices in $C_i \cup R_i$ for 2k + 2 < i < l.
- ▶ Lemma 12 (\star). Rule 4 is safe.

2.3 Reducing the teeth

▶ Lemma 13. Let (G = (V, E), k) be a yes-instance of TRIVIALLY PERFECT EDITING, and (C, R) be a comb of G such that there exist $a, b \in \{1, \ldots, l\}$ with $\sum_{a \leq i \leq l} |R_i| \geq 2k + 1$ and $\sum_{b \leq i < a} |R_i| \geq 2k + 1$. Then there exists an optimal k-edition F of G such that for every $m \in \{1, \ldots, b-1\}$, the vertices of R_m are all adjacent to the same vertices of $V(G) \setminus R_m$ in $G \triangle F$, and F contains no pair of vertices of R_m .

Proof. Let F be an optimal k-edition of G and $H = G \triangle F$. There exist $v_2 \in (R_a \cup R_{a+1} \cup \cdots \cup R_l)$ and $v_1 \in (R_b \cup R_{b+1} \cup \cdots \cup R_{a-1})$ unaffected by F. The neighborhood of v_1 in $H \setminus R$ must be a clique: indeed, if there exist $x, y \in N_G(v_1) \setminus R$ such that $xy \notin E(H)$, then since $(N_G(v_1) \setminus R) \subseteq (N_G(v_2) \setminus R)$, the vertices $\{v_1, x, v_2, y\}$ would induce a C_4 . Let $1 \le m < b$, we will construct an edition F_m such that $|F_m| \le |F|$, F_m contains no pair of vertices included in R_m and the vertices of R_m are all adjacent to the same vertices in $G \triangle F_m$. Applying this construction iteratively to each R_m , $1 \le m < b$ will yield an edition F^* that verifies the desired properties.

Let S be a maximal clique in H that contains $N_G(v_1)\backslash R$ and v_1 , and let K_1, \ldots, K_r be the connected components of $H\backslash S$. Observe that K_1, \ldots, K_r respect the conditions i, ii and iii of Lemma 6 with S. Let $v_m \in R_m$ be a vertex contained in the least number of pairs of F with the other element in S.

Denote by N the set of vertices of S adjacent to v_m in graph H. Let H' be the graph constructed from $H \setminus R_m$ and $G[R_m]$ by adding the edges $N \times R_m$, and F_m be the edition such that $H' = G \triangle F_m$. By construction $|F_m| \le |F|$, we will now show that H' is trivially perfect.

We can observe that $R_m \cap S = \emptyset$ (because v_1 is unaffected by F and is non-adjacent with R_m in G) and therefore that S is a maximal clique of $H \setminus R_m$.

By construction of H', S is also a maximal clique of H' and R_m is a connected component of $H_m \setminus S$. Let $K'_1, \ldots, K'_{r'}$ be the connected components of $(H \setminus R_m) \setminus S$. Sets $K'_1, \ldots, K'_{r'}$ verify the conditions i, ii and iii of Lemma 6 with respect to S in $H \setminus R_m$ and thus also in H'. Moreover $H'[S \cup R_m]$ is trivially perfect and $(N_{H'}(R_m) \times R_m) \subseteq E(H')$ by construction. The family $\bigcup_{1 \leq i \leq r} \{N_H(K_i)\}$ is nested according to Lemma 6, and, by construction of H', $\bigcup_{1 \leq i \leq r'} \{N_{H'}(K'_i)\} \subseteq \bigcup_{1 \leq i \leq r} \{N_H(K_i)\}$. We also have that $N \in \bigcup_{1 \leq i \leq r} \{N_H(K_i)\}$. Indeed, let $K(v_m)$ the connected component of $H \setminus S$ containing v_m , according to condition iii from Lemma 6 we have $N_H(K(v_m)) = N_H(v_m) \cap S = N$. Therefore the family $\bigcup_{1 \leq i \leq r} \{N_{H'}(K'_i)\} \cup \{N\}$ is also nested. By Lemma 6 applied on H' and S, graph H' is trivially perfect.

As mentioned previously, we can apply this construction iteratively to each R_m , $1 \le m < b$ and obtain an edition F^* that verifies the desired properties.

- ▶ Rule 5. Given a comb (C,R) of G such that there exist $a,b \in \{1,\ldots,l\}$ with $\Sigma_{a \leq i \leq l} |R_i| \geq 2k+1$ and $\Sigma_{b \leq i < a} |R_i| \geq 2k+1$. Then for every $i \in \{1,\ldots,b-1\}$, replace R_i by a clique of size $min(|R_i|,k+1)$ with the same neighborhood.
- ▶ Lemma 14 (\star). Rule 5 is safe.
- ▶ Lemma 15 (*). Let (G = (V, E), k) be an instance of TRIVIALLY PERFECT EDITING such that Rules 2 to 5 are not applicable. Then, for every comb (C, R) of G, $|C \cup R| = O(k^2)$.
- ▶ **Lemma 16** (*). Given an instance (G = (V, E), k) of TRIVIALLY PERFECT EDITING, Rules 4 and 5 can be exhaustively applied in polynomial time.

3 Bounding the size of a reduced instance

We now prove thoroughly that any reduced yes-instance of TRIVIALLY PERFECT EDITING contains $O(k^3)$ vertices. To that end, we need the following definition and result.

- ▶ Definition 17 (LCA-closure [16]). Let T = (V, E) be a rooted tree and $A \subseteq V(T)$. The lowest common ancestor-closure (LCA-closure) A' of A is obtained as follows. Initially, set A' = A. Then, as long as there exist $x, y \in A'$ whose lowest common ancestor w is not in A', add w to A'. The LCA-closure of A is the last set A' obtained using this process.
- ▶ Lemma 18 ([16]). Let T = (V, E) be a rooted tree, $A \subseteq V(T)$ and A' = LCA-closure(A). Then $|A'| \leq 2 \cdot |A|$ and for every connected component C of $T \setminus A'$, $|N_T(C)| \leq 2$.
- ▶ **Theorem 19.** TRIVIALLY PERFECT EDITING admits a kernel with $O(k^3)$ vertices.

Proof. Let (G = (V, E), k) be a reduced yes-instance of Trivially Perfect Editing and F a k-edition of G. Let $H = G \triangle F$ and $\mathcal{T} = (T, \mathcal{B})$ the universal clique decomposition of H. The graph G is not necessarily connected, thus T is a forest. Let A be the set of nodes $t \in V(T)$ such that the bag B_t contains a vertex affected by F. Since $|F| \leq k$, we have $|A| \leq 2k$. Let $A' \subseteq V(T)$ be the set containing the nodes of LCA-closure(A) and the root of each connected component of T (in case the closure does not contain them). According to Lemma 18 and Rule 1 which implies that there are at most 2k connected components in G and thus 2k roots, we have |A'| < 6k.

Let D be a connected component of $T \setminus A'$. We can observe that, by construction of A' (which for every pair of nodes, contains also the smallest common ancestor in T), only three cases are possibles (see Figure 4):

- $N_T(D) = \emptyset$ (D is a connected component of T).
- $N_T(D) = \{a\}$ (D is a subtree of T whose parent is $a \in A'$).
- $N_T(D) = \{a_1, a_2\}$ with one of the nodes $a_1, a_2 \in A'$ being an ancestor of the other in T. We will say that these connected components are respectively of type 0, 1 or 2. For $D \subseteq V(T)$, we denote by $W(D) = \bigcup_{t \in D} B_t$ the set of vertices of G corresponding to bags of D.

There is no connected component of type 0 or else W(D) would be a connected component of G inducing a trivially perfect graph. Rule 1 would have been applied to this component, contradicting the fact that G is a reduced instance.

Now consider the set of type 1 components D_1, D_2, \ldots, D_r of $T \setminus A'$ attached in T to the same node $a \in A'$. We show that $W_a = W(D_1) \cup W(D_2) \cup \cdots \cup W(D_r)$ is a trivially perfect module of G. In the graph H, W_a is by construction a module of the decomposition. Since

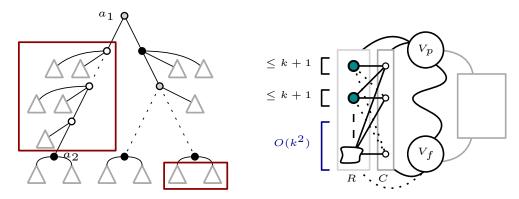


Figure 4 (Left) universal clique decomposition of a connected component of *H*. (Right) shape of a reduced comb.

no vertex of W_a was affected by the edition F, W_a is also a module of G, trivially perfect by heredity. By Lemma 8, we have $|W_a| = O(k^2)$. There are at most $|A'| \leq 6k$ such sets W_a , thus the set of vertices of G in bags of type 1 components is of size $O(k^3)$.

Now consider the type 2 connected components D of $T \setminus A'$ which have two neighbors in T. Let a_1 and a_2 be these neighbors, one being the ancestor of the other, say a_1 is the ancestor of a_2 . Let t_1, \ldots, t_l be the nodes of the tree on the path from a_1 to a_2 , in this order. The component D can be seen as a comb of shaft $(B_{t_1}, \ldots, B_{t_l})$. More precisely, by construction of the universal clique decomposition, W(D) can be partitioned into a comb (C, R) of H: the critical clique decomposition of C is $(C_1 = B_{t_1}, \ldots, C_l = B_{t_l})$, and each R_i corresponds to the union of bags of the subtrees rooted at t_i which do not contain t_{i+1} , for $1 \leq i < l$, and to the union of bags of the subtrees rooted at t_l which do not contain a_2 , for i = l. Since (C, R) was not affected by F, it is also a comb of G. Thus for each type 2 component D, W(D) contains $O(k^2)$ vertices by Lemma 15. Since T is a forest, it can contain at most $|A'| - 1 \leq 6k - 1$ such components in $T \setminus A$. Therefore the set of bags containing type 2 connected components of $T \setminus A$ contains $O(k^3)$ vertices.

It remains to bound the set of vertices of G which are in bags of A'. The vertices corresponding to nodes of $A' \setminus A$ are critical cliques of G, and are hence of size at most k+1 by Rule 2. Thus the set of vertices in bags of $A' \setminus A$ is of size $O(k^2)$. The vertices corresponding to nodes of A are critical cliques in H but not necessarily of G. Let B_a be a bag corresponding to a node $a \in A$. We will show that B_a is covered by at most 2k+1 critical cliques of G, which by Rule 2 will imply that B_a contains $O(k^2)$ vertices of G, and thus the set of vertices in bags of A' is of size $O(k^3)$.

To see this, observe that B_a is a critical clique of H, and that G is obtained from H by editing at most k pairs of vertices. A result from [31] claims that, starting from a graph H and editing an edge, we add at most two critical cliques. The same arguments allow to claim that if B is a set of vertices covered by at most p critical cliques in H, and if H' is obtained by editing a pair of vertices x, y of H, then p+2 critical cliques are enough to cover B in H'. To be complete, we now show this claim. Let $C_1, C_2, \ldots, C_p, \ldots, C_q$ be the critical cliques of H, suppose that B is covered by the first p cliques C_1, \ldots, C_p . For each $i, 1 \leq i \leq q$, the set $C_i'' = C_i \setminus \{x,y\}$ is a clique module (not necessarily maximal) of H'. In particular, each C_i'' is contained in a critical clique C_i' of H (the C_i' are not necessarily distinct). Let C'(x) and C'(y) be the critical cliques of H' containing respectively x and y. Clearly, the critical cliques $C_1', \ldots, C_p', C'(x)$ and C'(y) of H' cover the vertices of B, showing our claim. By applying this argument k times (one for each pair of F) to the bag B_a , which was a critical clique of H, we conclude that it is covered by at most 2k+1 critical cliques of G. Thus $|B_a| = O(k^2)$ by Rule 2.

45:12 A Cubic Vertex-Kernel for Trivially Perfect Editing

We conclude that $|V(G)| = O(k^3)$. Finally, we claim that a reduced instance can be computed in polynomial time. Indeed, Lemma 7 states that it is possible to reduce exhaustively a graph under Rules 2 to 3. Once this is done, it remains to apply exhaustively Rules 4 and 5 which is ensured by Lemma 16.

4 Kernels for trivially perfect completion/deletion

In this section we show that the rules used for TRIVIALLY PERFECT EDITING are safe for TRIVIALLY PERFECT COMPLETION and TRIVIALLY PERFECT DELETION. First Rules 1, 2 and 3 are safe for both problems. Indeed, the safeness of Rule 2 directly follows from Lemma 1 and Rule 3 was shown safe in [13].

We will now argue that Rules 4 and 5 are also safe. Lemma 11 states that no trivially perfect edition for an instance (G = (V, E), k) of Trivially Perfect Editing affects a comb of G of length at least 2k + 2. This is also true when allowing only edge addition or edge deletion, implying the safeness of Rule 4 in both cases. In the proof of Lemma 13, for a trivially perfect edition F we construct another edition $F' \subseteq F$. In case F consists only of edge additions or deletions, it is also the case for F', thus Lemma 13 holds for Trivially Perfect Completion and Trivially Perfect Deletion and Rule 5 is safe for these problems.

The proof for the size of the kernel is the same as the proof of Theorem 19. Altogether, we obtain the following result.

▶ Theorem 20. TRIVIALLY PERFECT COMPLETION and TRIVIALLY PERFECT DELETION admits a kernel with $O(k^3)$ vertices.

5 Conclusion

We have provided a kernelization algorithm for Trivially Perfect Editing, producing a cubic vertex-kernel, hence improving upon the $O(k^7)$ -size kernel of [13]. The techniques extend to the deletion and completion versions of the problem, within the same bounds. A natural question is whether the size of the kernel for Trivially Perfect Editing can still be reduced – note that for Trivially Perfect Completion, Bathie et al. claim a quadratic kernel [2].

Some ideas used in this work remind of very similar techniques applied to kernelization problems for edge editing towards classes of graphs \mathcal{G} having a tree-like decomposition. The simplest case – like here or for the class of so-called 3-leaf power graphs, see [3] – is when the vertices of the graph can be partitioned into bags inducing modules, and these bags can be structured as nodes of a forest T, with specific adjacency rules. If an arbitrary graph G can be turned into a graph of class \mathcal{G} by editing at most k pairs of vertices, the edited pairs are in some set A of at most 2k bags. Again by taking the lowest common ancestor closure A' of A, set A' is of size O(k) and its removal from forest T will produce some chunks attached in T to 0, 1 or 2 nodes of A' (e.g., in [3], the authors speak of 1 and 2-branches, playing similar roles to modules and combs in this article). Kernelization algorithms can be obtained if we are able to reduce the bags themselves as well as the chunks, which hopefully have good structural properties. It is natural to wonder how general are these techniques, especially on subclasses of chordal graphs.

- References

- N. R. Aravind, R. B. Sandeep, and Naveen Sivadasan. Dichotomy results on the hardness of h-free edge modification problems. SIAM J. Discret. Math., 31(1):542-561, 2017. doi: 10.1137/16M1055797.
- 2 Gabriel Bathie, Nicolas Bousquet, and Théo Pierron. (Sub)linear kernels for edge modification problems towards structured graph classes. In preparation, 2021.
- 3 Stéphane Bessy, Christophe Paul, and Anthony Perez. Polynomial kernels for 3-leaf power graph modification problems. *Discret. Appl. Math.*, 158(16):1732–1744, 2010. doi:10.1016/j.dam.2010.07.002.
- 4 Stéphane Bessy and Anthony Perez. Polynomial kernels for proper interval completion and related problems. *Inf. Comput.*, 231:89–108, 2013. doi:10.1016/j.ic.2013.08.006.
- 5 Ivan Bliznets, Fedor V. Fomin, Marcin Pilipczuk, and Michał Pilipczuk. A subexponential parameterized algorithm for proper interval completion. SIAM J. Discret. Math., 29(4):1961–1987, 2015. doi:10.1137/140988565.
- 6 Pablo Burzyn, Flavia Bonomo, and Guillermo Durán. Np-completeness results for edge modification problems. *Discrete Applied Mathematics*, 154(13):1824–1844, 2006. doi:10.1016/j.dam.2006.03.031.
- 7 Leizhen Cai. Fixed-parameter tractability of graph modification problems for hereditary properties. *Inf. Process. Lett.*, 58(4):171–176, 1996. doi:10.1016/0020-0190(96)00050-6.
- 8 Leizhen Cai and Yufei Cai. Incompressibility of H-free edge modification problems. Algorithmica, 71(3):731-757, 2015. doi:10.1007/s00453-014-9937-x.
- 9 Yixin Cao and Yuping Ke. Improved kernels for edge modification problems, 2021. arXiv: 2104.14510.
- Yixin Cao and Dániel Marx. Chordal editing is fixed-parameter tractable. *Algorithmica*, 75(1):118–137, 2016. doi:10.1007/s00453-015-0014-x.
- Christophe Crespelle, Pål Grønås Drange, Fedor V. Fomin, and Petr A. Golovach. A survey of parameterized algorithms and the complexity of edge modification. *CoRR*, abs/2001.06867, 2020. arXiv:2001.06867.
- 12 Pål Grønås Drange, Fedor V. Fomin, Michał Pilipczuk, and Yngve Villanger. Exploring the subexponential complexity of completion problems. *ACM Trans. Comput. Theory*, 7(4):14:1–14:38, 2015. doi:10.1145/2799640.
- Pål Grønås Drange and Michał Pilipczuk. A polynomial kernel for trivially perfect editing. Algorithmica, 80(12):3481–3524, 2018. doi:10.1007/s00453-017-0401-6.
- Maël Dumas, Anthony Perez, and Ioan Todinca. A cubic vertex-kernel for trivially perfect editing. *CoRR*, abs/2105.08549, 2021. arXiv:2105.08549.
- Ehab S El-Mallah and Charles J Colbourn. The complexity of some edge deletion problems. *IEEE transactions on circuits and systems*, 35(3):354–362, 1988. doi:10.1109/31.1748.
- Fedor V. Fomin, Daniel Lokshtanov, Neeldhara Misra, and Saket Saurabh. Planar F-deletion: Approximation, kernelization and optimal FPT algorithms. In 53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012, pages 470–479. IEEE Computer Society, 2012. doi:10.1109/FOCS.2012.62.
- 17 Fedor V Fomin, Daniel Lokshtanov, Saket Saurabh, and Meirav Zehavi. *Kernelization: theory of parameterized preprocessing*. Cambridge University Press, 2019.
- Fedor V. Fomin and Yngve Villanger. Subexponential parameterized algorithm for minimum fill-in. SIAM J. Comput., 42(6):2197–2216, 2013. doi:10.1137/11085390X.
- Esha Ghosh, Sudeshna Kolay, Mrinal Kumar, Pranabendu Misra, Fahad Panolan, Ashutosh Rai, and M. S. Ramanujan. Faster parameterized algorithms for deletion to split graphs. Algorithmica, 71(4):989–1006, 2015. doi:10.1007/s00453-013-9837-5.
- 20 Martin Charles Golumbic. Trivially perfect graphs. Discret. Math., 24(1):105-107, 1978. doi:10.1016/0012-365X(78)90178-4.

45:14 A Cubic Vertex-Kernel for Trivially Perfect Editing

- 21 Martin Charles Golumbic, Haim Kaplan, and Ron Shamir. On the complexity of DNA physical mapping. Advances in Applied Mathematics, 15(3):251-261, 1994. doi:10.1006/aama.1994. 1009.
- Sylvain Guillemot, Frédéric Havet, Christophe Paul, and Anthony Perez. On the (non-)existence of polynomial kernels for P_l-free edge modification problems. Algorithmica, 65(4):900–926, 2013. doi:10.1007/s00453-012-9619-5.
- 23 Jiong Guo. Problem kernels for NP-complete edge deletion problems: Split and related graphs. In Takeshi Tokuyama, editor, Algorithms and Computation, 18th International Symposium, ISAAC 2007, Sendai, Japan, December 17-19, 2007, Proceedings, volume 4835 of Lecture Notes in Computer Science, pages 915–926. Springer, 2007. doi:10.1007/978-3-540-77120-3_79.
- Pavol Hell, Ron Shamir, and Roded Sharan. A fully dynamic algorithm for recognizing and representing proper interval graphs. *SIAM J. Comput.*, 31(1):289–305, 2001. doi: 10.1137/S0097539700372216.
- 25 Haim Kaplan, Ron Shamir, and Robert Endre Tarjan. Tractability of parameterized completion problems on chordal, strongly chordal, and proper interval graphs. SIAM J. Comput., 28(5):1906–1922, 1999. doi:10.1137/S0097539796303044.
- Stefan Kratsch and Magnus Wahlström. Two edge modification problems without polynomial kernels. In Jianer Chen and Fedor V. Fomin, editors, Parameterized and Exact Computation, 4th International Workshop, IWPEC 2009, Copenhagen, Denmark, September 10-11, 2009, Revised Selected Papers, volume 5917 of Lecture Notes in Computer Science, pages 264–275. Springer, 2009. doi:10.1007/978-3-642-11269-0_22.
- Yunlong Liu, Jianxin Wang, and Jiong Guo. An overview of kernelization algorithms for graph modification problems. *Tsinghua Science and Technology*, 19(4):346–357, 2014. doi: 10.1109/TST.2014.6867517.
- Dániel Marx and R. B. Sandeep. Incompressibility of h-free edge modification problems: Towards a dichotomy. In Fabrizio Grandoni, Grzegorz Herman, and Peter Sanders, editors, 28th Annual European Symposium on Algorithms, ESA 2020, September 7-9, 2020, Pisa, Italy (Virtual Conference), volume 173 of LIPIcs, pages 72:1-72:25. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.ESA.2020.72.
- James Nastos and Yong Gao. Familial groups in social networks. Soc. Networks, 35(3):439-450, 2013. doi:10.1016/j.socnet.2013.05.001.
- Jaroslav Nesetril and Patrice Ossona de Mendez. On low tree-depth decompositions. *Graphs Comb.*, 31(6):1941–1963, 2015. doi:10.1007/s00373-015-1569-7.
- Fábio Protti, Maise Dantas da Silva, and Jayme Luiz Szwarcfiter. Applying modular decomposition to parameterized cluster editing problems. *Theory Comput. Syst.*, 44(1):91–104, 2009. doi:10.1007/s00224-007-9032-7.
- 32 Jing-Ho Yan, Jer-Jeong Chen, and Gerard J Chang. Quasi-threshold graphs. *Discrete applied mathematics*, 69(3):247–255, 1996.
- 33 Mihalis Yannakakis. Computing the minimum fill-in is NP-complete. SIAM Journal on Algebraic Discrete Methods, 2(1):77–79, 1981. doi:10.1137/0602010.