On the hardness of eliminating small induced subgraphs by contracting edges

Daniel Lokshtanov[⋆], Neeldhara Misra[†], and Saket Saurabh[⋆,◇]

[⋆]University of Bergen, Norway
[†]Indian Institute of Science, Bangalore
[◇]Institute of Mathematical Sciences, India
neeldhara@csa.iisc.ernet.in, daniello.ii.uib.no, saket@imsc.res.in

Abstract

Graph modification problems such as vertex deletion, edge deletion or edge contractions are a fundamental class of optimization problems. Recently, the parameterized complexity of the CONTRACTIBILITY problem has been pursued for various specific classes of graphs. Usually, several graph modification questions of the deletion variety can be seen to be FPT if the graph class we want to delete into can be characterized by a finite number of forbidden subgraphs. For example, to check if there exists k vertices/edges whose removal makes the graph C_4-free, we could simply branch over all cycles of length four in the given graph, leading to a search tree with $O(4^k)$ leaves. Somewhat surprisingly, we show that the corresponding question in the context of contractibility is in fact W[2]-hard. An immediate consequence of our reductions is that it is W[2]-hard to determine if at most k edges can be contracted to modify the given graph into a chordal graph. More precisely, we obtain following results:

- C_ℓ-FREE CONTRACTION is W[2]-hard if and only if $\ell \geqslant 4$.
- P_ℓ-FREE CONTRACTION is W[2]-hard if and only if $\ell \geqslant 3$.

We believe that this opens up an interesting line of work in understanding the complexity of contractibility from the perspective of the graph classes that we are modifying into.

1 Introduction

Graph modification problems constitute a broad and fundamental class of graph optimization problems. Typically, we are interested in knowing if a given input graph G is "close enough" to a graph H or a graph in a class of graphs $\mathcal{H}$. In the latter case, the goal is usually to see if G can be easily morphed into a graph with a certain property, and the class $\mathcal{H}$ is used to describe the said property [3]. Some of the most prevalent notions of closeness are defined in terms of vertex or edge deletion, or edge contraction. For example, when defined in terms of vertex deletion, one might ask if at most k vertices can be deleted to make the graph edgeless (here we are modifying into the class of empty graphs), and this is the classic VERTEX COVER problem.

In this work, we will restrict ourselves to the context of contractibility questions, and in particular, we would be contracting into graph classes that are described in terms of their induced forbidden subgraphs. In a $\mathcal{H}$-CONTRACTIBILITY problem, given a graph G and a positive integer k, the objective is to check if there exists a subset of at most k edges which, if contracted, lead to a graph in $\mathcal{H}$. Such questions are usually NP-complete on general graphs, and have recently received a lot of attention in the context of parameterized complexity. For example, it is known that the BIPARTITE CONTRACTION problem is FPT, and this is the contraction analog of EDGE BIPARTIZATION, which is the fundamental and well-studied question of whether k edges can be removed to make a given graph bipartite [9, 6]. This result involved an interesting combination of techniques, including iterative compression, important separators, and irrelevant vertices. Also, the problems of determining if k edges can be contracted to obtain a tree, or a path, are known to be FPT using a non-trivial application of color coding [7]. The PLANAR CONTRACTION problem was also shown to be FPT recently [5], again using irrelevant vertex techniques combined with an application of Courcelle's theorem.

Questions of contractibility have been investigated quite extensively when the input graph is restricted to being chordal, usually yielding polynomial time algorithms (see, for instance, [8, 2]). However, the natural question of CHORDAL CONTRACTION, while known to be NP-complete [1], remains un-investigated in the parameterized context. Before considering algorithms for CHORDAL CONTRACTION, we first explored the apparently easier question of contracting edges to obtain a C_4-free graph, that is, a graph with no induced cycles of length four. Notice that the vertex-deletion analog of this question is almost trivial from a parameterized point of view: we could simply branch over all cycles of length four in the given graph, leading to a search tree with $O(4^k)$ leaves. This is true of most problems which require us to "hit" a constant number of constant-sized forbidden subgraphs using a constrained budget. However, when we ask the same question in the context of contraction, the scenario is dramatically different: it is no longer true that a copy of a forbidden object can only be destroyed by edges that form the object — rather, edges contracted from "outside" the copy could also contribute towards its elimination. Therefore, the number of choices for branching is no longer obviously bounded. In fact, we find that the C_4-FREE CONTRACTION question turns out to be W[2]-hard, which we find rather surprising, considering the finite nature of the forbidden subgraph characterization of the graph class that we are interested in contracting to.

It turns out that our reduction also implies the hardness of CHORDAL CONTRACTION. On a closely related note, we show that the P_i-FREE CONTRACTION problem is also W[2]-hard. On the positive side, we show that it is FPT to determine if k edges can be contracted so that the resulting graph is a complete graph. In this case, the forbidden subgraph is just a single non-edge or an induced path on two edges. Further, we remark that it is easily checked that K_i-FREE CONTRACTION is FPT by the search tree technique. In this case, since the forbidden object, being a complete graph, cannot be "destroyed from outside", the branching is exhaustive.

The reason for describing the graph class $\mathcal{H}$ in terms of its forbidden subgraphs is to open up questions regarding a general characterization of the parameterized complexity of the problem in terms of the forbidden subgraphs, possibly analogous to the theorem of Asano and Hirata [1]. In this work, our goal is to motivate and initiate a study in this direction, by providing somewhat unexpected answers to a few specific cases.

Our Contributions. Let $\mathcal{H}$ be a graph class that has a forbidden induced subgraph characterization, and let $\mathcal{F}$ be the forbidden induced subgraphs for $\mathcal{H}$. Then, the $\mathcal{H}$ CONTRACTION question, or equivalently the $\mathcal{F}$-FREE CONTRACTION problem, is the following.

> $\mathcal{F}$-FREE CONTRACTION **Parameter:** k
> **Input:** A graph $G = (V, E)$ and a positive integer k
> **Question:** Is there a subset of at most k edges such that G/F has no induced copies of graphs $H \in \mathcal{F}$?

The C_ℓ-FREE CONTRACTION problem is known to be NP-complete. [1] for all fixed integer $\ell \geqslant 3$. We show, by a simple reduction from the HITTING SET problem, that the C_ℓ-FREE CONTRACTION problem is W[2]-hard for $\ell \geqslant 4$. Consequently, we establish that CHORDAL CONTRACTION is W[2]-hard. Further, we show that P_γ-FREE CONTRACTION is W[2]-hard for all $\gamma \geqslant 3$, while contracting to K_i-free graphs (for $i \geqslant 3$) and cliques turn out to be FPT.

The paper is organized as follows. After introducing some notation and preliminary notions in Section 2, we turn to the reductions. We first show that the C_4-FREE CONTRACTION problem is W[2]-hard, and subsequently describe a generalization. This is followed by the reduction for P_γ-FREE CONTRACTION.

We conclude with the tractable cases and suggestions for future directions.

2 Preliminaries

In this section we state some basic definitions related to parameterized complexity and graph theory, and give an overview of the notation used in this paper. Our notation for graph theoretic notions is standard and follows Diestel [4]. We summarize some of the frequently used concepts here. For a finite set V, a pair $G = (V, E)$ such that $E \subseteq V^2$ is a graph on V. The elements of V are called *vertices*, while pairs of vertices (u, v) such that $(u, v) \in E$ are called *edges*. We also use $V(G)$ and $E(G)$ to denote the vertex set and the edge set of G, respectively. In the following, let $G = (V, E)$ and $G' = (V', E')$ be graphs, and $U \subseteq V$ some subset of vertices of G. Let G' be a subgraph of G. If E' contains all the edges $\{u, v\} \in E$ with $u, v \in V'$, then G' is an *induced subgraph* of G, *induced by* V', denoted by $G[V']$. For any $U \subseteq V$, $G \setminus U = G[V \setminus U]$. For $v \in V$, $N_G(v) = \{u \mid (u, v) \in E\}$.

The *contraction* of edge xy in G removes vertices x and y from G, and replaces them by a new vertex, which is made adjacent to precisely those vertices that were adjacent to at least one of the vertices x and y. A graph G is *contractible* to a graph H, or H-*contractible*, if H can be obtained from G by a sequence of edge contractions. Equivalently, G is H-contractible if there is a surjection $\varphi : V(G) \to V(H)$, with $W(h) = \{v \in V(G) \mid \varphi(v) = h\}$ for every$h \in V(H)$, that satisfies the following three conditions: **(1)** for every $h \in V(H)$, $W(h)$ is a connected set in G; **(2)** for every pair $h_i, h_j \in V(H)$, there is an edge in G between a vertex of $W(h_i)$ and a vertex of $W(h_j)$ if and only if $h_i h_j \in E(H)$; **(3)** $\mathcal{W} = \{W(h) \mid h \in V(H)\}$ is a partition of $V(G)$. We say that $\mathcal{W}$ is an H-*witness structure* of G, and the sets $W(h)$, for $h \in V(H)$, are called *witness sets* of $\mathcal{W}$. It is easy to see that if we contract every edge $uv \in E(G)$, such that u and v belong to the same witness set, then we obtain a graph isomorphic to H. Hence G is H-contractible if and only if it has an H-witness structure.

A path is a sequence of vertices $v_1, v_2, \ldots, v_r$ such that $(v_i, v_i + 1) \in E$ for all $1 \leqslant i \leqslant r - 1$. A cycle is a sequence of vertices $v_1, v_2, \ldots, v_r$ such that $(v_i, v_i + 1) \in E$ for all $1 \leqslant i \leqslant r - 1$, and $(v_r, v_1) \in E$. A graph is said to be *chordal*, or *triangulated* if it has no induced cycles of length four or more.

Parameterized Complexity. A parameterized problem is denoted by a pair $(Q, k) \subseteq \Sigma^* \times \mathbb{N}$. The first component Q is a classical language, and the number k is called the parameter. Such a problem is *fixed–parameter tractable* (FPT) if there exists an algorithm that decides it in time $O(f(k)n^{O(1)})$ on instances of size n. Next we define the notion of parameterized reduction.

Definition 1. *Let A, B be parameterized problems. We say that A is (uniformly many:1)* ***fpt-reducible*** *to B if there exist functions $f, g : \mathbb{N} \to \mathbb{N}$, a constant $\alpha \in \mathbb{N}$ and an algorithm Φ which transforms an instance (x, k) of A into an instance $(x', g(k))$ of B in time $f(k)|x|^\alpha$ so that $(x, k) \in A$ if and only if $(x', g(k)) \in B$.*

A parameterized problem is considered unlikely to be fixed-parameter tractable if it is $W[i]$-hard for some $i \geqslant 1$. To show that a problem is $W[2]$-hard, it is enough to give a parameterized reduction from a known $W[2]$-hard problem. Throughout this paper we follow this recipe to show a problem $W[2]$-hard.

3 Hardness of Contraction Problems

In this section we address the parameterized complexity of C_j-free contraction, chordal contraction and P_j-free contraction. All the reductions are from the Hitting Set problem, and have a similar underlying flavor. We would begin by creating a separate induced instance of a forbidden object for every set in the universe. Then we will typically have edges corresponding to the elements in the universe, and the edges are placed to ensure that contracting them will “kill” exactly those forbidden objects that correspond to the sets that the element belongs to. Often, this is achieved with the following wireframe: we anchor all the edges corresponding to vertices of the universe to a common vertex, and let the forbidden object “dangle” from the same vertex. Now, to encode the instance, we add edges between the free end of the edges that correspond to the vertices of the universe and a suitably chosen vertex

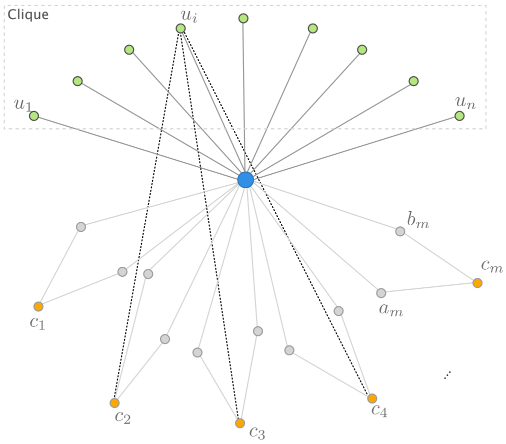

Figure 1: The construction for reducing to C_4-free graphs. In this example, the adjacencies corresponding to the HITTING SET are illustrated for the element x_i, which is assumed to belong to the sets S_2, S_3 and S_4.

of the relevant forbidden objects. We would expect that this generic idea is realized in different ways depending on what the forbidden objects are. In the rest of this section, we will describe two instances of specific reductions in detail, formalizing the ideas described above.

3.1 Contracting to C_ℓ-free Graphs

Our first exploration is to do with the problem of contracting to graphs that contain no induced cycles of length ℓ. In the interest of exposition, we begin by explaining the reduction for the case of reducing to C_4-free graphs. Since it turns out that the reduced instance has no longer induced cycles, this reduction already implies the hardness of contracting k edges to obtain a chordal graph. We will subsequently describe an easy generalization of the construction.

C_4-FREE CONTRACTION	**Parameter:** k
Input: A graph $G = (V, E)$ and a positive integer k	
Question: Is there a subset of at most k edges such that G/F has no induced cycles of length four?	

We reduce from the HITTING SET problem. Let $(U, \mathcal{F})$ be an instance of HITTING SET, where $U = \{x_1, x_2, \ldots, x_n\}$ and $\mathcal{F} = \{S_1, S_2, \ldots, S_m\}$, where each $S_i \subseteq U$. We denote the reduced instance to be constructed by $G = (V, E)$. The vertex set consists of a special central vertex, denoted by g, one vertex for each element $x_i \in U$, denoted by u_i, and three vertices for every set S_i in the family $\mathcal{F}$, denoted by a_i, b_i, c_i. We now describe the edges. The central vertex is adjacent to every vertex other than $\{c_i \mid 1 \leqslant i \leqslant m\}$. We impose a clique on the vertices that correspond to elements of the universe. Next, we add the edges $(a_i c_i)$ and $(b_i c_i)$ for every $1 \leqslant i \leqslant m$. Finally, for every $x_i \in S_j$, we add the edge (u_i, c_j). This completes the construction. Formally, the instance is given as follows (also see Figure 1).
$V := \{g\} \cup \{u_i \mid 1 \leqslant i \leqslant n\} \cup \left(\bigcup_{1 \leqslant i \leqslant m} \{a_i, b_i, c_i\}\right)$ and

$$E := \left(\bigcup_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m} \{(g, u_i), (g, a_j), (g, b_j)\} \right) \cup \left(\bigcup_{1 \leqslant j \leqslant m} \{(c_j, a_j), (c_j, b_j)\} \right)$$
$$\cup \{(u_i, u_j) \mid 1 \leqslant i \neq j \leqslant n\} \cup \{(u_i, c_j) \mid 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m, \text{ and } x_i \in S_j\}$$

We begin by identifying the induced cycles of length four in the graph G. This will help us in showing the correctness of the reduction.

Proposition 1. *The only induced cycles of length four in the graph* G *are formed by the vertex sets given below:*

- $\{g, a_i, c_i, b_i\}$*, for all* $1 \leqslant i \leqslant m$*,*
- $\{u_i, g, a_j, c_j\}$*, for all* $x_i \in S_j$*, and*
- $\{u_i, g, b_j, c_j\}$*, for all* $x_i \in S_j$*.*

Proof. Clearly, for all $1 \leqslant i \leqslant m$, the vertices $\{g, a_i, b_i, c_i\}$ induce a four-cycle, and for all $x_i \in S_j$, the vertices $\{u_i, g, t, c_j\}$ (where t is either a_j or b_j) induce a four-cycle as well. Assume, for the sake of contradiction, that there exists an induced four-cycle other than the ones accounted for, with the vertex set $C := \{w, x, y, z\}$. Let T denote the vertex subset $\{g\} \cup \{u_i \mid 1 \leqslant i \leqslant n\}$. Note that $|C \cap T| \leqslant 2$, since $G[T]$ is a clique, and $G[C]$ is an induced cycle of length four. Notice that $G \setminus T$ is acyclic, so C intersects T in either one or two vertices.

First, consider the case when $|T \cap C| = 1$, and without loss of generality, let $T \cap C = \{w\}$. Suppose $w \neq g$. Then $w = u_i$ for some $1 \leqslant i \leqslant n$. Notice that u_i is adjacent to vertices $N_i := \{c_j \mid x_i \in S_j\}$. However, it is easily checked that no two vertices in N_i share a common neighbor in $G \setminus T$. Indeed, for $1 \leqslant p \neq q \leqslant m$, $N_{G \setminus T}(c_p) = \{b_p, a_p\}$ and $N_{G \setminus T}(c_q) = \{b_q, a_q\}$. Therefore, $N(x) \cap N(y) \cap G \setminus T = \emptyset$ for all $x, y \in N_i$, and w cannot be extended to an induced four-cycle from vertices in $G \setminus T$. On the other hand, let $w = g$. Then, let the neighbors of w in the four-cycle C are x and z. Clearly, $x := a_j$ or $x := b_j$, for some $1 \leqslant j \leqslant m$. Without loss of generality, let $x := a_j$. Now, $z \neq b_j$, since in this case, the unique cycle that w, x and z can be completed to is already accounted for. Thus, $z := v_\ell^{(a)}$ or $z := v_\ell^{(b)}$ for some $\ell \neq j$. Again, in this case, z and x share no common neighbors in $G \setminus T$, and we are done.

The second case is when $|T \cap C| = 2$. Again, without loss of generality, let $T \cap C = \{w, x\}$. First, consider the situation when $w \neq g$ and $x \neq g$. Let $w = u_p$ and $x = u_q$. For w and x to be part of an induced four-cycle, w and x need to have private neighbors in $G \setminus T$ that are adjacent. However, it is easy to verify that $N(u_p) \cup N(u_q)$ in $G \setminus T$ is an independent set. Therefore, there is no way of extending this choice of w and x to a four-cycle. Finally, suppose $w = g$, and let $x = u_p$. Every neighbor of u_p is c_i for some i and every neighbor of g lies in $\{a_j, b_j \mid 1 \leqslant j \leqslant m\}$. The only possibilities for forming induced four-cycles arise from choosing $c_i \in N(u_p)$ and either a_j or b_j with $j = i$. However, note that all of these cycles have been accounted for in the statement of the proposition. This completes the proof. □

We now turn to the correctness of the reduction.

Lemma 1. *The graph* G *described as above is a* YES*-instance of* C_4-FREE CONTRACTION *if, and only if,* $(U, \mathcal{F})$ *is a* YES*-instance of* HITTING SET*.*

Proof. First, suppose $(U, \mathcal{F})$ is a YES-instance of HITTING SET, and let $S \subseteq U$ be a solution. Consider the edges corresponding to S in G, that is, let F be defined as $\{(g, u_i) \mid$ for all $u_i \in S$. We claim that G/F has no induced cycles of length four. Clearly, the proposed solution has the appropriate size, since we are picking one edge corresponding to every element of the hitting set, which is assumed to have size at most k. We now argue that the suggested set indeed forms a solution. First, notice that when the edge (g, u_i) is contracted, g becomes adjacent to every c_j for which $x_i \in S_j$ (see Figure 2). Since we are contracting vertices that form a hitting set, notice that for every $1 \leqslant j \leqslant m$, the edge (g, c_j) is present in G/F. By Proposition 1, the only induced four-cycles that need to be killed are as follows:

- $\{g, a_i, c_i, b_i\}$, for all $1 \leqslant i \leqslant m$,
- $\{u_i, g, a_j, c_j\}$, for all $x_i \in S_j$, and
- $\{u_i, g, b_j, c_j\}$, for all $x_i \in S_j$.

Notice that the edge (g, c_j) is a chord with respect to all these cycles, and this completes the argument in the forward direction.

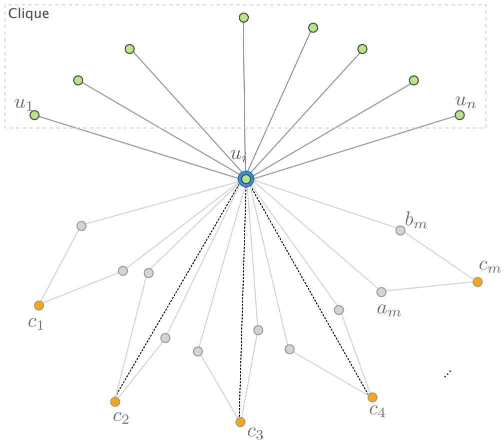

Figure 2: This figure illustrates what happens when the edge (g, u_i) is contracted. As shown in the figure, all the induced cycles of length four that were created by vertices c_j for $u_i \in S_j$ are now destroyed.

In the reverse direction, suppose we have a subset of k edges, say F, such that G/F has no induced cycles of length four. We first argue that there exists a solution F that does not use any edge from the C_4 corresponding to the sets. Suppose F contains an edge e that is of the form (g, a_j) or (g, b_j). Clearly, contracting such an edge only affects the cycle $\{g, a_j, c_j, b_j\}$. Let x_i be any element of S_j. Consider the set $F^\star$ given by $F \setminus \{e\} \cup \{(g, u_i)\}$. It is easy to see that $F^\star$ is also a solution, since $G/F^\star$ has a chord in the cycle $\{g, a_j, c_j, b_j\}$. A similar argument shows that if F contains an edge of the form (a_j, c_j) or (b_j, c_j), then it can be replaced with an appropriately chosen edge of the form (g, u_i).

Finally, if F contains an edge e of the form (u_i, c_j), then notice that the only four-cycles of G that become triangulated in $G/\{e\}$ are: $\{g, a_j, c_j, b_j\}$, $\{u_i, g, a_j, c_j\}$, and $\{u_i, g, b_j, c_j\}$. All of these cycles also become triangulated when the edge (u_i, g) is contracted instead. Therefore, in this case also, we note that the set $F^\star$ given by $F \setminus \{e\} \cup \{(g, u_i)\}$ is also a solution.

Let $T^\star$ denote the set $\{u_1, \ldots, u_n\}$. By above arguments we have shown that there exists a solution F that is contained in the clique formed on $T^\star \cup \{g\}$. We are now ready to describe a hitting set S of size at most k. Let $\mathcal{W}$ be a G/F-witness structure of G and let $W(g)$ be the witness set that contains the global vertex g. Observe that since $G[W(g)]$ is connected we have that the $|W(g)| \leqslant k+1$. We take S as $W(g) \setminus S$. Clearly, the size of S is at most k. It is also straightforward to see that S forms a hitting set. Indeed, consider any set $S_j \in \mathcal{F}$. Now consider the four-cycle given by $\{g, a_j, c_j, b_j\}$. Since it is triangulated, it must be the case that there is a $x_i \in S_i$ for which $u_i \in W(g)$, and hence $x_i \in S$. This concludes the reverse direction of the reduction. ☐

From Lemma 2, and the hardness of the HITTING SET problem, we have the following:

Theorem 1. *The* C_4-FREE CONTRACTION *problem is* W[2]-hard *when parameterized by the size of the solution.*

Notice that in the analysis of Proposition 1, it is evident that the graph has no induced cycles of length five or more. Therefore, exactly the same arguments can be used to derive the fact that the problem of CHORDAL CONTRACTION, where we ask if k edges can be contracted to make the input graph chordal, is W[2]-hard when parameterized by k.

Corollary 1. *The* CHORDAL CONTRACTION *problem is* W[2]-hard *when parameterized by the size of the solution.*

Now we consider the C_ℓ-FREE CONTRACTION problem for $i \geqslant 5$. Notice that if we replace the cycles of length four with cycles of length ℓ in the reduction above, and make the vertices in u_i adjacent to the $\lfloor(\ell/2)\rfloor^{\text{th}}$ vertex in the cycle, then our claims follow by very similar arguments. We describe the construction and because of the similarity of the arguments defer the details of the correctness to the full version of this paper.

As before, let $(U, \mathcal{F})$ be an instance of HITTING SET, where $U = \{x_1, x_2, \ldots, x_n\}$ and $\mathcal{F} = \{S_1, S_2, \ldots, S_m\}$, where each $S_i \subseteq U$. We denote the reduced instance to be constructed by $G = (V, E)$. The vertex set consists of a special central vertex, denoted by g, one vertex for each element $x_i \in U$, denoted by u_i, and $(\ell - 1)$ vertices for every set S_i in the family $\mathcal{F}$, denoted by $a_i^1, a_i^2, \ldots, a_i^{\ell-1}$.

We now describe the edges. The central vertex is adjacent to the vertices u_i and a_j^1, $a_j^{\ell-1}$, for $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$. We impose a clique on the vertices that correspond to elements of the universe. Next, we add the edges (g, a_i^1), $(g, a_i^{\ell-1})$ and (a_i^j, a_i^{j+1}) for every $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant \ell - 2$. Finally, for every $x_i \in S_j$, we add the edge $(u_i, a_j^{\lfloor \ell/2 \rfloor})$. This completes the construction.

The proof of correctness is along the same lines as for the case of C_4-free contraction. In fact, for values of $\ell \geqslant 6$, there will be exactly m induced cycles of length ℓ in the graph G, as the cycles that use g, u_i and half of a cycle formed by a-vertices will not be of the requisite length, so the case analysis for the analog of Proposition 1 only simplifies. The detailed arguments are deferred to avoid repetition. This discussion brings us to the following theorem.

Theorem 2. *The* C_ℓ-FREE CONTRACTION *problem, for all fixed integer* $\ell \geqslant 4$, *is* W[2]-hard *when parameterized by the size of the solution.*

3.2 Contracting to P_γ-free Graphs

For the purposes of our discussion in this section, a path of length γ is a path on γ edges and $(\gamma + 1)$ vertices. For the problem of contracting to graphs that have no induced paths of length γ or longer, it turns out that the general case is easier to present, and we explain where the proof differs in the case of contracting to P_3-free graphs separately. For the case when $\gamma = 2$, in the next section, we show a FPT algorithm via an exponential kernel.

P_γ-FREE CONTRACTION	**Parameter:** k
Input: A graph $G = (V, E)$ and a positive integer k	
Question: Is there a subset of at most k edges such that G/F has no induced paths of length γ?	

Again, we reduce from HITTING SET. Let $(U, \mathcal{F})$ be an instance of HITTING SET, where $U = \{x_1, x_2, \ldots, x_n\}$ and $\mathcal{F} = \{S_1, S_2, \ldots, S_m\}$, where each $S_i \subseteq U$. We denote the reduced instance to be constructed by $G = (V, E)$. The vertex set consists of a special central vertex, denoted by g, one vertex for each element $x_i \in U$, denoted by u_i, and j vertices for every set S_i in the family $\mathcal{F}$, denoted by $a_i^1, a_i^2, \ldots, a_i^{\lfloor(\gamma/2)\rfloor}, b_i^1, b_i^2, \ldots, b_i^{\lceil(\gamma/2)\rceil}$. For readability, whenever it is clear from the context, we use ℓ to denote $(\gamma/2)$, appropriately rounded up or down. Also, let $T = \{u_1, \ldots, u_n\}$ denote the subset of vertices corresponding to the elements of the universe, and for every $1 \leqslant i \leqslant m$, denote the sets $\{a_i^1, a_i^2, \ldots, a_i^\ell\}$ $\{b_i^1, b_i^2, \ldots, b_i^\ell\}$ by A_i and B_i, respectively.

We now describe the edges. To begin with, we impose a clique on $T \cup \{g\}$. Next, add edges to ensure that the sets A_i and B_i induce paths of lengths $\ell - 1$ and $\ell + 1$, starting at a_i^1 and b_i^1, respectively. Further, we make the central vertex g and *all* the vertices in T adjacent to a_i^1 and b_i^1 for all $1 \leqslant i \leqslant m$. Notice that there is now an induced path of length γ starting at a_ℓ, going via g and ending at b_ℓ. To encode the hitting set structure, for every $x_i \in S_j$, make u_i adjacent to a_j^ℓ, a_j^2, b_j^ℓ, and b_j^2.

A formal summary of the construction is below, also see Figure 3. Here,

$$V := \{g\} \cup \{u_i \mid 1 \leqslant i \leqslant n\} \cup \left(\bigcup_{1 \leqslant i \leqslant \ell} \{a_j^i \mid 1 \leqslant j \leqslant m\} \right) \cup \left(\bigcup_{1 \leqslant i \leqslant \ell} \{b_j^i \mid 1 \leqslant j \leqslant m\} \right),$$

and the edge set

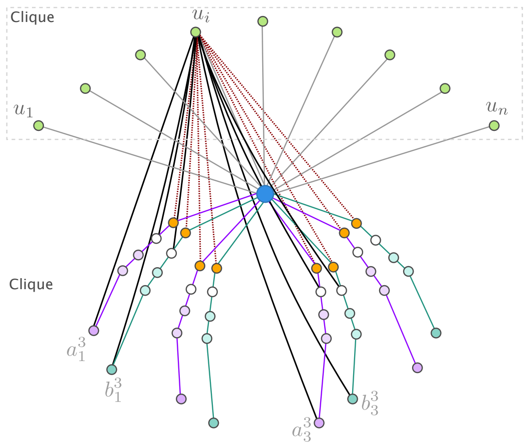

Figure 3: The construction for reducing to P_j-free graphs. In this example, $\gamma = 10$, and the adjacencies corresponding to the Hitting Set are illustrated for the element x_i, which is shown as belonging to the sets S_1 and S_3.

$$\begin{aligned} E \; := \; & \{(u_i, x) \mid 1 \leqslant i \leqslant n, x \in T \cup \{g\}\} \\ & \cup \Big(\bigcup_{1 \leqslant i \leqslant m} \{\{(g, a_i^1,)\} \cup \{(a_i^j, a_i^{j+1}) \mid 1 \leqslant j \leqslant \ell - 1\} \Big) \\ & \cup \Big(\bigcup_{1 \leqslant i \leqslant m} \{\{(g, b_i^1,)\} \cup \{(b_i^j, b_i^{j+1}) \mid 1 \leqslant j \leqslant \ell - 1\} \Big) \\ & \cup \{(u_i, a_j^\ell), (u_i, b_j^\ell)(u_i, a_j^2), (u_i, b_j^2) \mid 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m, \text{ and } x_i \in S_j\} \\ & \cup \{(u_i, a_j^1), (u_i, b_j^1) \mid 1 \leqslant i \leqslant n \text{ and } 1 \leqslant j \leqslant m\}. \end{aligned}$$

We now make a structural observation analogous to Proposition 1 in the previous reduction. Here, we show that the only induced paths with γ or more edges in the graph are those that correspond to the paths involving A_i, B_i and the central vertex g, for all $1 \leqslant i \leqslant m$.

Proposition 2. [1] *If $\gamma \geqslant 4$, then the only induced paths of length γ or more in the graph G described in Section 3.2 are those that are formed by the sequences:*

$$x_i^\ell, x_i^{\ell-1}, \ldots, x_i^2, x_i^1, g, y_j^1, y_j^2, \ldots, y_j^{\ell-1}, y_j^\ell,$$

for any $1 \leqslant i, j \leqslant m$, and $x \in \{a, b\}$ and $y \in \{a, b\}$.

Proof. Clearly, for every $1 \leqslant i, j \leqslant m$, there is an induced path given by:

$$x_i^\ell, \ldots, x_i^2, x_i^1, g, y_j^1, y_j^2, \ldots, y_j^\ell,$$

where $x \in \{a, b\}$ and $y \in \{a, b\}$. The lengths of these paths are either γ (when γ is even), or between $\gamma - 1$ and $\gamma + 1$, depending on the choice of ℓ.

We now show that there are no other induced paths of length γ or more in G. For the rest of this discussion, we use ℓ to refer to $\lceil (\gamma/2) \rceil$. Note that as a result of this convention, we have $2\ell = (\gamma + 1)$. Consider any induced path P_v starting at a vertex $v \in G$, and let the path be given by the sequence $v, t_1, t_2, \ldots, t_r$. We let T denote the subset of vertices corresponding to the elements of the universe. Further, let $A := \{a_i^\ell \mid 1 \leqslant i \leqslant m\}$, and $B := \{b_i^\ell \mid 1 \leqslant i \leqslant m\}$.

[1] Proofs for results marked with $\star$ are moved to appendix.

Case 1: $\nu = g$. First, suppose $t_1 \in T$, and without loss of generality, let $t_1 := u_i$ Then $t_2 \notin T$, since g is adjacent to every vertex in T. Therefore, $t_2 \in N(u_i) \setminus T$, in other words, $t_2 \in A \cup B$. Without loss of generality, let $t_2 := a_j^\ell$ for some suitable j. From here, the path cannot come back to T, since g is adjacent to every vertex in T. The other case is when t_3 is $a_j^{\ell-1}$. From here, the path can only extend further via the sequence a_j^q, $q = \{\ell-1, \ell-2 \ldots, 3\}$ (recall that a_j^2 is adjacent to u_i whenever a_j^ℓ is adjacent to u_i, by construction). The path we have finally arrived at is therefore the following:

$$g, u_i, a_j^\ell, \ldots, a_j^4, a_j^3.$$

Notice that the length of the path thus derived is at most $2 + (\ell - 3) = \ell - 1 \leqslant (\gamma - 1)$.

The other case is when $t_1 \notin T$, in this case t_1 is either a_i^1 or b_i^1 for some $1 \leqslant i \leqslant n$. Notice that other than moving to a_i^2 or b_i^2 (respectively), the only other option is to come back to T, which is not feasible since g is adjacent to every vertex in T. Therefore the next vertex in the path is uniquely determined as either a_i^2 or b_i^2. The subsequent vertices have degree two in G, and repeating this argument, it is clear that the choice of the induced path starting with the edge (g, t_1) is uniquely determined to be:

$$g, a_j^1, a_j^2, \ldots, a_j^\ell.$$

if $t_1 := a_i^1$, and is

$$g, b_j^1, b_j^2, \ldots, b_j^\ell.$$

if $t_1 := b_i^1$. Notice that the only neighbors of a_j^ℓ and b_j^ℓ that are not already used in the path are in T, and g is adjacent to every vertex in T, therefore the described paths are maximal. The maximum length of either of these paths is at most $\ell < \gamma$, and we are done with the case when $\nu = g$.

Case 2: $\nu \in T$. Without loss of generality, let $\nu := u_i$. Note that $N(\nu) = \{g\} \cup T \cup X$, where $X \subseteq A \cup B$. We do a case analysis on the choice of $t_1 \in N(\nu)$.

If $t_1 = g$, then $t_2 \notin T$ and the other options either a_i^1 or b_i^1 for some $1 \leqslant i \leqslant n$. However, both of these vertices are adjacent to u_i, and therefore the path is truncated at length one on this route.

Next, suppose $t_1 \in X$. Without loss of generality, let $t_1 := a_j^\ell$. From here, $t_2 \notin T$, since u_i is adjacent to every vertex in T. Therefore, the path is determined by the sequence:

$$a_j^\ell, \ldots, a_j^3.$$

Note that the only way to extend the path further would be with the vertex a_j^2, but this is adjacent to u_i by construction. Therefore, we have a path of length $1 + (\ell - 3)$, which is at most γ, as desired.

Finally, suppose $t_1 \in T$. Then $t_2 \notin T \cup \{g\}$ and therefore, it must be the case that $t_2 \in (A \cup B)$. As discussed above, the longest path from here can only be of length $(\ell - 3)$, so the total length of the path is $2 + (\ell - 3)$, which is also at most γ.

Case 3: $\nu \neq g$, **and** $\nu \in G \setminus T$. In this case, we let $\nu := a_j^d$ for some $1 \leqslant j \leqslant m$, and $1 \leqslant d \leqslant \ell - 1$. We can fix j without loss of generality, and the argument for the case when $\nu := b_j^d$ is symmetric. We now argue based on the cases when $d = \ell$, $d = 1$, $d = 2$ and $3 \leqslant d \leqslant \ell$. The arguments are somewhat tedious but straightforward.

The case when $d = \ell$. Note that $N(\nu) = \{a_j^{\ell-1}\} \cup X$, where $X \subseteq T$. Again, we do a case analysis based on the choices of t_1 from $N(\nu)$. First, suppose $t_1 := a_j^{\ell-1}$. Now, the path has no choice until it reaches a_j^2. At this point, it can move to a_j^1 or it can move to T.

- Suppose the path moves to T. The longest induced path starting at a vertex in T has length at most $(\ell-1)$, and we have used $\ell-2$ edges in arriving at a_j^2, and one additional edge to move from a_j^2 to T. Thus, the total length of the path we have is at most $(\ell-1)+(\ell-1)$, therefore, this is a path of length at most $\gamma-1$.
- On the other hand, suppose the path moves to a_j^1. From here, we distinguish two further cases depending on whether the next vertex, say w is g or in T.
  - $w \in T$. Notice that since the path contains a_j^1, we cannot move from w to either g, or any other vertex in T, since a_j^1 is adjacent to all such vertices. The only other option is to move to a suitable vertex in $A \cup B$, say a_x^ℓ. From here, since we cannot return to T, the only option is to follow the indices path along the $a's$, namely, $a_x^{\ell-1}, \ldots, a_x^3$. Notice that the length of the last leg of the path is $(\ell-3)$ and this is maximal because w is adjacent to a_x^2. Therefore, the total length of the path is $(\ell-1)+2+(\ell-3) = 2\ell-2 \leqslant (\gamma+1)-2 = \gamma-1$.
  - $w = g$. Having come to g, if we move to some $a_{j'}^1$, or $b_{j'}^1$, then the only way to complete such a path will be of the kind that we have accounted for in the statement of the proposition. However, we cannot move to any other vertex from g because the only other neighbors are in T, and every vertex in T is adjacent to a_j^1, which is already on the path.

The case when $3 \leqslant d \leqslant \ell-1$. We have $v := a_j^d$. If the next vertex on the path, that is t_1, happens to be a_j^{d-1}, then the arguments made in the first case (when $d = \ell$) apply and we have the desired conclusion. On the other hand, suppose $t_1 := a_j^{d+1}$. Then the first part of the path is determined until we reach a_j^ℓ, that is, we have the initial vertices of the path given by:

$$a_j^d, a_j^{d+1}, \ldots, a_j^\ell$$

Now from here, the path must move to T. The longest induced path starting at a vertex in T has length at most $(\ell-1)$, and we have used $\ell-3$ edges in arriving at a_j^ℓ, and one additional edge to move from a_j^ℓ to T. Thus, the total length of the path we have is at most $(\ell-2)+(\ell-1)$, therefore, this is a path of length at most $\gamma-2$.

The case when $d = 2$. Here, again, we have two cases — we either move from a_j^2 to a_j^1 or to a_j^3. If we move to a_j^1, then the arguments are identical to those we used in the first case, when $v := a_j^\ell$. Since the longest path starting from a_j^3 is at most $(\gamma-2)$, the total length of the path in the latter case is $(\gamma-1)$.

The case when $d = 1$. Starting from a_j^1, we can move to either a vertex in T, g, or a_j^2. If we move to either T or g, then the arguments are identical to those we used in the first case, when $v := a_j^\ell$. On the other hand, if we move to a_j^2, then we cannot move back to T, having started from a_j^1. Therefore, we must move along the path $\{a_j^2, a_j^3, \ldots, a_j^\ell\}$, and at this point we have a maximal path of length ℓ, since the only further choices lead us back to T. We are done since $\ell < \gamma$.

With this we exhaust all the cases towards the claim in the proposition and conclude the argument. □

We now turn to the correctness of the reduction.

Lemma 2. *Let γ be a fixed integer $\geqslant 4$. The graph G described as above is a* YES*-instance of* P_γ-FREE CONTRACTION *if, and only if,* $(U, \mathcal{F})$ *is a* YES*-instance of* HITTING SET.

Proof. First, suppose $(U, \mathcal{F})$ is a YES-instance of HITTING SET, and let $S \subseteq U$ be a solution. Consider the edges corresponding to S in G, that is, let F be defined as $\{(g, u_i) \mid \text{ for all } u_i \in S\}$. We claim that G/F has no induced paths of length γ. Clearly, the proposed solution has the appropriate size, since we are picking one edge corresponding to every element of the hitting set, which is assumed to have size at most k. We now argue that the suggested set indeed forms a solution. First, notice that when the edge (g, u_i) is contracted, g becomes adjacent to every a_j^2, a_j^ℓ, b_j^2 and b_j^ℓ for which $x_i \in S_j$ (see also Figure 4). Since we are contracting vertices that form a hitting set, notice that for every $1 \leqslant j \leqslant m$, the edges (g, a_j^ℓ) and (g, b_j^ℓ) is present in G/F. By Proposition 2, the only induced paths that need to be killed are as follows: $a_i^\ell, a_i^{\ell-1}, \ldots, a_i^2, a_i^1, g, b_j^1, b_j^2, \ldots, b_j^{\ell-1}, b_j^\ell$, for any $1 \leqslant i, j \leqslant m$ (the other possibilities are

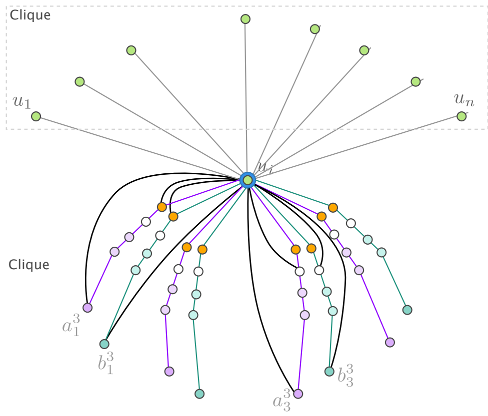

Figure 4: This figure illustrates what happens when the edge (g, u_i) is contracted. Notice that all the relevant induced paths are destroyed.

symmetric). Notice that the edges (g, a_j^ℓ) and (g, b_j^ℓ) are chord with respect to paths $a_j^\ell, a_j^{\ell-1}, \ldots, a_j^2, a_j^1, g$ and $g, b_j^1, b_j^2, \ldots, b_j^{\ell-1}, b_j^\ell$ respectively, and this completes the argument in the forward direction.

In the reverse direction, suppose we have a subset of k edges, say F, such that G/F has no induced paths of length γ. We first argue that there exists a solution F that does not use any edge from the paths $A_j = a_j^\ell, a_j^{\ell-1}, \ldots, a_j^2, a_j^1, g$ and $B_j = g, b_j^1, b_j^2, \ldots, b_j^{\ell-1}, b_j^\ell$ for any $1 \leqslant j \leqslant m$. Suppose F contains an edge e that is of the form (x, y) from the path A_j (argument for B_j is symmetric). Clearly, contracting such an edge only affects the induced paths that use A_j as a subpath. Let x_i be any element of S_j. Consider the set $F^\star$ given by $F \setminus \{e\} \cup \{(g, u_i)\}$. It is easy to see that $F^\star$ is also a solution, since $G/F^\star$ has a chord in the path A_j.

Finally, if F contains an edge e of the form (u_i, x) where (u_i, x) is on the path A_j (the case where (u_i, x) is on the path B_j is symmetric), then notice that the only induced paths of G that get destroyed are those involving A_j. All of these induced paths can also be destroyed when the edge (u_i, g) is contracted instead. Therefore, in this case also, we note that the set $F^\star$ given by $F \setminus \{e\} \cup \{(g, u_i)\}$ is also a solution.

Let T denote the set $\{u_1, \ldots, u_n\}$. By above arguments we have shown that there exists a solution F that is contained in the clique formed on $T \cup \{g\}$. We are now ready to describe a hitting set S of size at most k. Let $\mathcal{W}$ be a G/F-witness structure of G and let $W(g)$ be the witness set that contains the global vertex g. Observe that since $G[W(g)]$ is connected we have that the $|W(g)| \leqslant k+1$. We take S as $W(g) \setminus S$. Clearly, $|S| \leqslant k$. It is also straightforward to see that S forms a hitting set. Indeed, consider any set $S_j \in \mathcal{F}$. Now consider the induced path given by $a_j^\ell, a_j^{\ell-1}, \ldots, a_j^2, a_j^1, g, b_j^1, b_j^2, \ldots, b_j^{\ell-1}, b_j^\ell$. Since this does not exist anymore, it must be the case that there is a $x_i \in S_i$ for which $u_i \in W(g)$, and hence $x_i \in S$. This concludes the reverse direction of the reduction. □

Notice that the lemma above holds for $\gamma \geqslant 4$, and the case when $\gamma = 3$ is shown to be tractable in the next section. The reason the case when $\gamma = 3$ is a degenerate case with respect to the above construction is the following. It turns out that it is possible to choose vertices $(a_i^\ell, u_p, u_q, b_j^\ell)$ appropriately so that this sequence corresponds to an induced P_3 that is different from the collection of paths that we have introduced by design. It is difficult to destroy such paths in the forward direction. However, the reduction can be modified slightly: namely, by imposing a clique of all the a_j^ℓ and b_j^ℓ vertices, except for the (a_j^ℓ, b_j^ℓ) edges. We will also need to make $(k+1)$ copies of every pair of vertices that we introduced for the sets of the family $\mathcal{F}$. With this revised construction, it is easily checked that there are no induced paths of length three other than the ones we want. The additional copies are merely introduced to ensure that there is always a solution that does not contract any of the newly introduced edges. The proof of correctness is very similar to the proof of Lemma 2 above. The complete details are deferred to a full version of this paper. To conclude, from Lemma 2, and the hardness of the HITTING SET problem, we have the following:

Theorem 3. *The* P_γ-FREE CONTRACTION *problem is* W[2]-hard *for all fixed integer* $\gamma \geqslant 3$ *when parameterized by the size of the solution.*

4 A Few Tractable Cases

In this section we give FPT algorithm for a few cases of $\mathcal{F}$-FREE CONTRACTION – namely K_ℓ-FREE CONTRACTION for every fixed integer $\ell \geqslant 3$ and P_2-FREE CONTRACTION. That is we prove the following.

Theorem 4. K_ℓ-FREE CONTRACTION. *For every fixed integer* $\ell \geqslant 3$ *and* P_2-FREE CONTRACTION *are* FPT*.*

Proof. To solve K_ℓ-FREE CONTRACTION we do as follows. Given an undirected graph G on n vertices and a positive integer k, we first find a clique K_ℓ and then iteratively contract every edge of this clique and recursively search for solution of size $k-1$ in the contracted graph. Since the forbidden object, being a complete graph, cannot be "destroyed from outside", the branching is exhaustive. This leads to a FPT algorithm with running time $O(\ell^{2k} n^{O(1)})$. Observe that this implies C_3-FREE CONTRACTION is FPT.

Now we show that P_2-FREE CONTRACTION is FPT. Let (G, k) be an instance to P_2-FREE CONTRACTION. For simplicity we assume that G is connected, else we could apply our algorithm to each connected component separately. It is well know that a graph does not have induced P_2 if and only if it is a clique. To solve the problem given (G, k), in polynomial time, we output an equivalent instance (G', k) with at most $O(4^k k)$ vertices. Given the small sized equivalent instance we can try all possible choice of at most k edges as possible solution and check whether their contraction leads to a clique.

Two vertices u and v are called *twins* if $N(u) = N(v)$. If there exists a set S of size at least $2k+1$ such that for all u and v in S we have that $N(u) = N(v)$ (that is, S is a set of twins of size at least $2k+1$) then delete an arbitrary vertex w from S. Observe that since we are only allowed to contract at most k edges, the number of vertices that can be adjacent to one of the contracted edges is upper bounded by $2k$. One can easily show using this observation that (G, k) is a yes instance of P_2-FREE CONTRACTION if and only of $(G \setminus \{w\}, k)$ is a yes instance of P_2-FREE CONTRACTION. We apply this twin reduction rule as long as possible. If (G, k) is a yes instance then there exists a set F of at most k edges whose contraction lead to a clique. Let W be the end points of edges in F. Clearly, $|W| \leqslant 2k$. Now we group the vertices of $G' \setminus W$ with their neighborhood in W. This implies that there are at most 4^k groups. Observe that vertices in the same group are twins. Thus, the size of each twin class is upper bounded by $2k$. This implies that if (G, k) is a yes instance and thus (G', k) is a yes instance then the number of vertices in G' is upper bounded by $4^k \cdot 2k$. Hence, if (G', k) has more than $4^k \cdot 2k$ vertices then we return that (G, k) is a no instance else (G', k) is the required small sized equivalent instance. This completes the proof. ☐

5 Conclusions

In this paper we initiated the study of $\mathcal{F}$-FREE CONTRACTION problem and answered questions when $\mathcal{F}$ consisted of a fixed cycle or a path of a particular length. Quite surprisingly we showed that the C_j-FREE CONTRACTION problem is W[2]-hard for $i \geqslant 4$ and thus CHORDAL CONTRACTION is W[2]-hard. We also showed that P_j-FREE CONTRACTION is W[2]-hard for $i \geqslant 3$, and established that P_2-FREE CONTRACTION is FPT.

An interesting, and potentially challenging, question would be to characterization the parameterized complexity of $\mathcal{F}$-FREE CONTRACTION in terms of properties of the forbidden subgraphs $\mathcal{F}$. On the other hand, it will also be interesting examine if there are subclasses of graphs on which the problems of C_j-FREE CONTRACTION (for $j \geqslant 4$) and P_j-FREE CONTRACTION (for $j \geqslant 3$) admit FPT algorithms while being NP-complete.

References

[1] Takao Asano and Tomio Hirata. Edge-contraction problems. *J. Comput. Syst. Sci.*, 26(2):197–208, 1983.

[2] Rémy Belmonte, Pinar Heggernes, and Pim van 't Hof. Edge contractions in subclasses of chordal graphs. *Discrete Applied Mathematics*, 160(7-8):999–1010, 2012.

[3] Leizhen Cai. Fixed-parameter tractability of graph modification problems for hereditary properties. *Inf. Process. Lett.*, 58(4):171–176, 1996.

[4] Reinhard Diestel. *Graph Theory.* Springer-Verlag, Heidelberg, third edition, 2005.

[5] Petr A. Golovach, Pim van 't Hof, and Daniël Paulusma. Obtaining planarity by contracting few edges. *Theor. Comput. Sci.*, 476:38–46, 2013.

[6] Sylvain Guillemot and Dániel Marx. A faster fpt algorithm for bipartite contraction. *CoRR*, abs/1305.2743, 2013.

[7] Pinar Heggernes, Pim van 't Hof, Benjamin Lévêque, Daniel Lokshtanov, and Christophe Paul. Contracting graphs to paths and trees. In *Parameterized and Exact Computation - 6th International Symposium, IPEC*, pages 55–66, 2011.

[8] Pinar Heggernes, Pim van 't Hof, Benjamin Lévêque, and Christophe Paul. Contracting chordal graphs and bipartite graphs to paths and trees. *Electronic Notes in Discrete Mathematics*, 37:87–92, 2011.

[9] Pinar Heggernes, Pim van 't Hof, Daniel Lokshtanov, and Christophe Paul. Obtaining a bipartite graph by contracting few edges. In *Foundations of Software Technology and Theoretical Computer Science, FSTTCS*, pages 217–228, 2011.