Polynomial Kernels for 3-Leaf Power Graph Modification Problems*

Stéphane Bessy[1], Christophe Paul[2], and Anthony Perez[1]

[1] LIRMM - Université Montpellier 2, France
[2] CNRS, LIRMM - Université Montpellier 2, France

Abstract. A graph $G = (V, E)$ is a 3-leaf power iff there exists a tree T the leaf set of which is V and such that $(u, v) \in E$ iff u and v are at distance at most 3 in T. The 3-leaf power edge modification problems, *i.e.* edition (also known as the Closest 3-Leaf Power), completion and edge-deletion are FPT when parameterized by the size of the edge set modification. However, a polynomial kernel was known for none of these three problems. For each of them, we provide a kernel with $O(k^3)$ vertices that can be computed in linear time. We thereby answer an open question first mentioned by Dom, Guo, Hüffner and Niedermeier [9].

1 Introduction

The combinatorial analysis of experimental data-sets naturally leads to graph modification problems. For example, extracting a threshold graph from a dissimilarity on a set is a classical technique used in clustering and data analysis to move from a numerical to a combinatorial data-set [1,17]. The edge set of the threshold graph aims at representing the pairs of elements which are close to each another. As the dissimilarity reflects some experimental measures, the edge set of the threshold graph may reflect some false positive or negative errors. So for the sake of cluster identification, the edge set of the threshold graph has to be edited in order to obtain a disjoint union of cliques. This problem, known as cluster editing, is fixed-parameter tractable (see e.g. [12,14,25]) and efficient parameterized algorithms have been proposed to solve biological instances with about 1000 vertices and several thousand edge modifications [2,6].

The (Proper) Interval Graph Completion problem is another example of graph modification problem which arises in the context of molecular biology for the dna physical mapping problem [13,18]. So, motivated by the identification of some hidden combinatorial structures on experimental data-sets, edge-modification problems cover a broad range of classical graph optimization problems, among which *completion* problems, *edition* problems and *edge-deletion* problems (see [21] for a recent survey). Though most of the edge-modification problems turn out to be NP-hard problems, efficient algorithms can be obtained to solve the natural parameterized version of some of them. Indeed, as long as the

* Work supported by the French research grant ANR-06-BLAN-0148-01 "Graph Decomposition and Algorithms - GRAAL".

J. Fiala, J. Kratochvíl, and M. Miller (Eds.): IWOCA 2009, LNCS 5874, pp. 72–82, 2009.

number k of errors generated by the experimental process is not too large, one can afford a time complexity exponential in k. A problem is *fixed parameterized tractable* (FPT for short) with respect to parameter k [11,22] whenever it can be solved in time $f(k) \cdot n^{O(1)}$, where $f(k)$ is an arbitrary computable function. The reduction to a polynomial kernel is one of the most efficient fixed parameterized algorithmic techniques. The idea is to preprocess the input in order to reduce its size while preserving the existence of a solution so that the size of the reduced instance is bounded by a polynomial in the parameter k. More formally, a problem is *kernelizable* if any instance (G, k) can be reduced in polynomial time (using *reduction rules*) into an instance (G', k') such that $k' \leqslant k$ and the size of G' is bounded by a function of k. Clearly having a kernel of small size is highly desirable and is an important issue in the context of applications [6,16].

This paper follows this line of research and studies the kernelization of edge-modification problems related to the family of *leaf powers*, graphs arising from a phylogenetic reconstruction context [19,20,23]. The goal is to extract, from a threshold graph G on a set S of species, a tree T, whose leaf set is S and such that the distance between two species is at most p in T iff they are adjacent in G (p being the value used to extract G from dissimilarity information). If such a tree T exists, then G is a *p-leaf power* and T is its *p-leaf root*. For $p \leqslant 5$, the p-leaf power recognition is polynomial time solvable [3,5], whereas the question is still open for p strictly larger than 5. Parameterized p-leaf power edge modification problems have been studied so far for $p \leqslant 4$. The edition problem for $p = 2$ is known as the Cluster Editing problem for which the kernel size bound has been successively improved in a series of recent papers [12,14,24], culminating in [15] with a kernel with $4k$ vertices. For larger values of p, the edition problem is known as the Closest p-Leaf Power problem. For $p = 3$ and 4, the Closest p-Leaf Power problem is known to be FPT [7,9], while its fixed parameterized tractability is still open for larger values of p. However, the existence of a polynomial kernel for $p \neq 2$ remained an open question [8,10]. Though the completion and edge-deletion problems are FPT for $p \leqslant 4$ [9,10], no polynomial kernel is known for $p \neq 2$ [15].

Our Results. We prove that the Closest 3-Leaf Power, the 3-Leaf Power Completion and the 3-Leaf Power Edge-Deletion problems admit a kernel with $O(k^3)$ vertices. We thereby answer positively to the open question of Dom, Guo, Hüffner and Niedermeier [9,10].

2 Preliminaries

The graphs we consider in this paper are undirected and loopless. The vertex set of a graph G is denoted by $V(G)$, with $|V(G)| = n$, and its edge set by $E(G)$, with $|E(G)| = m$. The *open* neighborhood of a vertex x is denoted by $N_G(x)$ and its *closed* neighborhood $N_G(x) \cup \{x\}$ by $N_G[x]$. We write $d_G(u, v)$ for sthe distance between two vertices u and v in G (in the notations, the reference to the graph G will be omitted when the context is clear). For a subset $S \subseteq V$, $d_S(u, v)$ denotes the distance between u and v within S. Two vertices x and y

of G are *true twins* if $N[x] = N[y]$. A *module* is a set of vertices S such that for any distinct vertices x and y of S, $N(x) \setminus S = N(y) \setminus S$. The subgraph induced by a subset S of vertices is denoted $G[S]$. If H is a subgraph of G, $G \setminus H$ stands for $G[V(G) \setminus V(H)]$. A graph family $\mathcal{F}$ is *hereditary* if for any graph $G \in \mathcal{F}$, any induced subgraph H of G also belongs to $\mathcal{F}$.

As the paper deals with undirected graphs, we abusively denote by $X \times Y$ the set of unordered pairs containing one element of X and one of Y. Let $G = (V, E)$ be a graph and F be a subset of $V \times V$, $G + F$ is the graph on vertex set V, the edge set of which is $E \vartriangle F$ (the symmetric difference between E and F). Such a set F is called an *edition* of G (we may also abusively say that $G + F$ is an edition). A vertex $v \in V$ is *affected* by an edition F whenever F contains an edge incident to v. Given a graph family $\mathcal{F}$ and given a graph $G = (V, E)$, a subset $F \subseteq V \times V$ is an *optimal $\mathcal{F}$-edition* of G if F is a set of minimum cardinality such that $G + F \in \mathcal{F}$. If we constrain F to be disjoint from E, we say that F is a *completion*, whereas if F is asked to be a subset of E, then F is an *edge deletion*. The problem we mainly consider is thus the following:

PARAMETERIZED CLOSEST 3-LEAF POWER:

Input : A graph $G = (V, E)$.
Parameter : $k \in \mathbb{N}$.
Question : Is there a 3-leaf power edition F of G such that $|F| \leqslant k$?

If we replace edition by deletion (resp.completion), we get the PARAMETERIZED 3-LEAF POWER EDGE-DELETION (resp. PARAMETERIZED 3-LEAF POWER COMPLETION) problem.

2.1 Critical Cliques

The notions of critical clique and critical clique graph, introduced in [20], have been recently successfully used in problems such as CLUSTER EDITING [15] and BICLUSTER EDITING [24].

Definition 1. *A* critical clique *of a graph G is a clique K which is a module and is maximal under this property.*

It follows from definition that the set $\mathcal{K}(G)$ of critical cliques of a graph G defines a partition of its vertex set V.

Definition 2. *Given a graph $G = (V, E)$, its* critical clique graph $\mathcal{C}(G)$ *has vertex set $\mathcal{K}(G)$ and edge set $E(\mathcal{C}(G))$ with*

$$(K, K') \in E(\mathcal{C}(G)) \Leftrightarrow \forall\ v \in K, v' \in K', (v, v') \in E(G)$$

The following lemma was used in the construction of polynomial kernels for CLUSTER EDITING and BICLUSTER EDITING problems in [24].

Lemma 1. *Let $G = (V, E)$ be a graph. If H is the graph $G + \{(u, v)\}$ with $(u, v) \in V \times V$, then $|\mathcal{K}(H)| \leqslant |\mathcal{K}(G)| + 4$.*

We now generalize a result used to obtain FPT algorithms for the Closest 3-Leaf Power problem [7]. A graph family $\mathcal{F}$ is said to be closed under *twin addition* if for any graph $G \in \mathcal{F}$, adding a twin of any of its vertices yields a graph of $\mathcal{F}$.

Lemma 2. *Let $\mathcal{F}$ be an hereditary graph family closed under true twin addition. For any graph $G = (V, E)$, there exists an optimal $\mathcal{F}$-edition (resp. $\mathcal{F}$-deletion, $\mathcal{F}$-completion) F such that any critical clique of $G + F$ is the disjoint union of a subset of critical cliques of G.*

In particular, this means that one can find an optimal solution that does not delete any edges within a critical clique. Furthermore, in this optimal solution, either all or no edges are added or deleted between two critical cliques. From now on, every considered optimal edition (resp. deletion, completion) is supposed to verify these two properties.

2.2 Leaf Powers

Definition 3. *Let T be an unrooted tree whose leaves are one-to-one mapped to the elements of a set V. The k-*leaf power *of T is the graph T^k, with $T^k = (V, E)$ where $E = \{(u, v) \mid u, v \in V$ and $d_T(u, v) \leqslant k\}$. We call T a k-*leaf root *of T^k.*

It is easy to see that for any k, the k-leaf power family of graphs satisfies the conditions of Lemma 2. In this paper we focus on the class of 3-leaf powers for which several characterizations are known, one of which propose a list of forbidden induced subgraphs [8]. The proofs of our kernel for the Closest 3-Leaf Power problem (or 3-Leaf Power Editing) rely on the well-known critical clique graph characterization and on a new one which is based on the join composition of graphs.

Theorem 1. *[7] A graph G is a 3-leaf power iff its critical clique graph $\mathcal{C}(G)$ is a forest.*

Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two disjoint graphs and let $S_1 \subseteq V_1$ and $S_2 \subseteq V_2$ be two non-empty subsets of vertices. The *join composition* of G_1 and G_2 on S_1 and S_2, denoted $(G_1, S_1) \otimes (G_2, S_2)$, results in the graph $H = (V_1 \cup V_2, E_1 \cup E_2 \cup (V(S_1) \times V(S_2)))$.

Theorem 2. *Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two connected 3-leaf powers. The graph $H = (G_1, S_1) \otimes (G_2, S_2)$, with $S_1 \subseteq V_1$ and $S_2 \subseteq V_2$, is a 3-leaf power if and only if one of the following conditions holds:*

1. *S_1 and S_2 are two cliques of G_1 and G_2 respectively, and if S_1 (resp. S_2) is not critical, then G_1 (resp. G_2) is a clique or,*
2. *there exists $v \in V_1$ such that $S_1 = N[v]$ and $S_2 = V_2$ is a clique.*

The following observation will help proving reduction rules.

Observation 3. *Let C be a critical clique of a 3-leaf power $G = (V, E)$. For any $S \subseteq V$, if the clique $C \setminus S$ is not critical in $G[V \setminus S]$, then the connected component of $G[V \setminus S]$ containing $C \setminus S$ is a clique.*

3 A Cubic Kernel for Closest 3-Leaf Power Problem

In this section, we present five preprocessing reduction rules the application of which leads to a kernel with $O(k^3)$ vertices for the Closest 3-Leaf Power problem. We first give two simple reduction rules; note that the second one was already used to obtain a kernel with $O(k^2)$ vertices for the parameterized Cluster Editing problem [24].

Rule 1. *If G has a connected component C such that $G[C]$ is 3-leaf power, then remove C from G.*

Rule 2. *If G has a critical clique K such that $|K| > k+1$, then remove $|K|-k-1$ vertices of K from $V(G)$.*

3.1 Branch Reduction Rules

We now assume that the input graph G is reduced under Rule 1 and Rule 2. The next three reduction rules use the fact that the critical clique graph of a 3-leaf power is a forest. The idea is to identify induced subgraphs of G, called *branches*, which correspond to subtrees of $\mathcal{C}(G)$. Such a subgraph is already a 3-leaf power.

Definition 4. *Let $G = (V, E)$ be a graph. An induced subgraph $G[S]$, with $S \subseteq V$, is a* branch *if S is the union of critical cliques $K_1, \ldots, K_r \in \mathcal{K}(G)$ such that the subgraph of $\mathcal{C}(G)$ induced by $\{K_1, \ldots, K_r\}$ is a tree.*

Let $B = G[S]$ be a branch of a graph G and let $K_1, \ldots, K_r$ be the critical cliques of G contained in S. We say that K_i ($1 \leqslant i \leqslant r$) is an *attachment point* of the branch B if it contains a vertex x such that $N_G(x)$ intersects $V(G) \setminus S$. A branch B is a *l-branch* if it has l attachment points. Our next three rules deal with 1-branches and 2-branches. In the following, we denote by B^R the subgraph of B in which the vertices of the attachment points have been removed. For an attachment point P of B, $N_B(P)$ denotes the set of neighbors of vertices of P in B.

Lemma 3. *Let $G = (V, E)$ be a graph and B be a 1-branch of G with attachment point P. There exists an optimal 3-leaf power edition F of G such that:*

1. *the set of affected vertices of B is a subset of $P \cup N_B(P)$ and*
2. *in $G + F$, the vertices of $N_B(P)$ are all adjacent to the same vertices of $V(G) \setminus V(B^R)$.*

Proof. Let F be an arbitrary optimal 3-leaf power edition of G. We construct from F another optimal 3-leaf power edition which satisfies the two conditions above. Let C be the critical clique of $H = G + F$ that contains P and set $C' = C \setminus B^R$. By Lemma 2, the set of critical cliques of G whose vertices belong to $N_B(P)$ contains two kind of cliques: $K_1, \ldots, K_c$, whose vertices are in C or adjacent to the vertices of C in H, and $K_{c+1}, \ldots, K_h$ whose vertices are not adjacent to the vertices of C in H. For $i \in \{1, \ldots, h\}$, let C_i be the connected component of B^R containing K_i.

Let us consider the three following induced subgraphs : G_1 the subgraph of G induced by the disjoint union of $C_1, \ldots, C_c$; G_2 the subgraph of G induced by the disjoint union of $C_{c+1}, \ldots, C_h$; and finally G', the subgraph of H induced by $V(G) \setminus V(B^R)$. Let us notice that these three graphs are 3-leaf powers. By Observation 3, if C' is not a critical clique of G', then the connected component of G' containing C' is a clique. Similarly, if K_i, for any $1 \leqslant i \leqslant c$, is not a critical clique of G_1, then the connected component of G_1 containing K_i is a clique. Thus, by Theorem 2, the disjoint union H' of G_2 and $(G', C') \otimes (G_1, \{K_1, \ldots, K_c\})$ is a 3-leaf power. By construction, the edge edition set F' such that $H' = G + F'$ is a subset of F and thus $|F'| \leqslant |F|$. Moreover, the vertices of B affected by F' all belong to $P \cup N_B(P)$, which proves the first point.

To state the second point, we focus on the relationship between the critical cliques K_i and C' in $H' = G + F'$. If some K_i is linked to C' in H' (i.e. $c > 1$), it means that the cost of adding the missing edges between K_i and C' (which, by Theorem 2, would also result in a 3-leaf power) is lower than the cost of removing the existing edges between K_i and C': $|K_i| \cdot |C' \setminus P| \leqslant |K_i| \cdot |P|$. On the other hand, if some K_j is not linked to C' in H' (i.e. $c < h$), we conclude that $|P| \leqslant |C' \setminus P|$. Finally, if both cases occur, we have $|P| = |C' \setminus P|$, and we can choose to add all or none of the edges between K_i and C'. In all cases, we provide an optimal edition of G into a 3-leaf power in which the vertices of $N_B(P)$ are all adjacent to the same vertices of $V(G) \setminus V(B^R)$. □

The safeness of the first 1-branch reduction rule follows from Lemma 3.

Rule 3. *If G contains a 1-branch B with attachment point P, then remove from G the vertices of B^R and add a new critical clique of size* $\min\{|N_B(P)|, k+1\}$ *adjacent to P.*

Our second 1-branch reduction rule considers the case where several 1-branches are attached to the rest of the graph by a join. The following lemma shows that

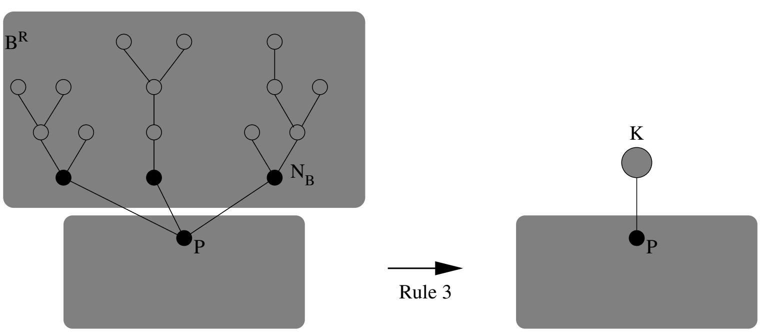

Fig. 1. On the left, a 1-branch B with attachment point P. On the right, the effect of Rule 3 which replaces B^R by a clique K of size $\min\{|N_B(P)|, k+1\}$.

under some cardinality conditions, the vertices of such 1-branches are not affected by an optimal 3-leaf power edition.

Lemma 4. *Let $G = (V, E)$ be a graph for which a 3-leaf power edition of size at most k exists. Let $B_1, \ldots, B_l$ be 1-branches, the attachment points $P_1, \ldots, P_l$ of which all have the same neighborhood N in $V \setminus \cup_{i=1}^{l} V(B_i)$. If $\sum_{i=1}^{l} |P_i| > 2k+1$, then there is no optimal 3-leaf power edition F of G that affects vertices of $\cup_{i=1}^{l} V(B_i)$.*

By Lemma 4, if there exists a 3-leaf power edition F of G such that $|F| \leqslant k$, then the 1-branches $B_1, \ldots, B_l$ can be safely replaced by two critical cliques of size $k+1$. This gives us the second 1-branch reduction rule.

Rule 4. *If G has several 1-branches $B_1, \ldots, B_l$, the attachment points $P_1, \ldots, P_l$ of which all have the same neighborhood N in $V \setminus \cup_{i=1}^{l} V(B_i)$ and if $\sum_{i=1}^{l} |P_i| > 2k+1$, then remove from G the vertices of $\cup_{i=1}^{l} V(B_i)$ and add two new critical cliques of size $k+1$ neighboring exactly N.*

3.2 The 2-Branch Reduction Rule

To complete the set of reduction rules, we need to consider 2-branches. So let B be a 2-branch with attachment points P_1 and P_2. The subgraph of G induced by the critical cliques of the unique path from P_1 to P_2 in $\mathcal{C}(B)$ is called the *main path* of B and denoted *path*(B). We say that B is *clean* if P_1 and P_2 are leaves of $\mathcal{C}(B)$, in which case we denote by Q_1 (resp. Q_2) the critical clique that neighbors P_1 (resp. P_2) in B.

Lemma 5. *Let B be a clean 2-branch of a graph $G = (V, E)$ with attachment points P_1 and P_2 such that path(B) contains at least 5 critical cliques. There exists an optimal 3-leaf power edition F of G which, if it affects vertices of B not in $V(P_1 \cup Q_1 \cup P_2 \cup Q_2)$, then it contains a min-cut of path(B).*

Rule 5. *If G has a clean 2-branch B such that path(B) is composed by at least 5 critical cliques, then remove from G all the vertices of $V(B)$ except those of $V(P_1 \cup Q_1 \cup P_2 \cup Q_2)$ and add four new critical cliques:*
- *K_1 (resp. K_2) of size $k+1$ adjacent to Q_1 (resp. Q_2);*
- *K_1' (resp K_2') adjacent to K_1 (resp. K_2) and such that K_1' and K_2' are adjacent and $|K_1'| \cdot |K_2'|$ equals the min-cut of path(B).*

3.3 Kernel Size and Time Complexity

Let us discuss the time complexity of the reduction rules. The 3-leaf power recognition problem can be solved in $O(n+m)$ time [4]. It follows that Rule 1 requires linear time. To implement the other reduction rules, we first need to compute the critical clique graph $\mathcal{C}(G)$, which, as noticed in [24], can be computed in linear time if we use modular decomposition algorithm (see [26] for a recent paper). Given $\mathcal{C}(G)$, which is linear in the size of G, it is easy to detect the critical

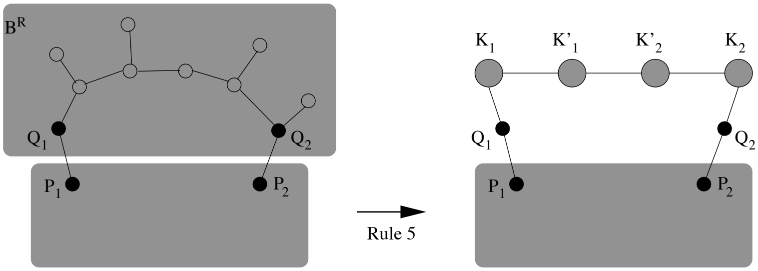

Fig. 2. A 2-branch B on the left (only pendant critical cliques are hanging on $path(B)$ since we can assume that the graph is reduced by the previous rules). On the right, the way Rule 5 reduces B.

cliques of size at least $k+1$. So, Rule 2 requires linear time. A search on $\mathcal{C}(G)$ can identify the 1-branches. It follows that the two 1-branches reduction rules (Rule 3 and Rule 4) can also be applied in $O(n+m)$ time. Let us now notice that in a graph reduced by the first four reduction rules, a 2-branch is a path to which pendant vertices are possibly attached. It follows that to detect a 2-branch B, such that $path(B)$ contains at least 5 critical cliques, we first prune the pendant vertices, and then identify in $\mathcal{C}(G)$ the paths containing only vertices of degree 2, and at least 5 of them. To do this, we compute the connected components of the graph induced on vertices of degree 2 in $\mathcal{C}(G)$. This shows that Rule 5 can be carried out in linear time.

Theorem 4. *The* Parameterized Closest 3-Leaf Power *problem admits a kernel with $O(k^3)$ vertices. Given a graph G, a reduced instance can be computed in linear time.*

Proof. The discussion above established the time complexity to compute a kernel. Let us determine the kernel size. Let $G = (V, E)$ be a reduced graph (*i.e.* none of the reduction rules applies to G) which can be edited into a 3-leaf power with a set $F \subseteq V \times V$ such that $|F| \leqslant k$. Let us denote $H = G + F$ the edited graph. For the sake of simplicity, we assume that H is connected, and thus $\mathcal{C}(H)$ is a tree. If $\mathcal{C}(H)$ is a forest, one has to apply the following arguments to each of its connected component, and then to sum up. We first show that $\mathcal{C}(H)$ has $O(k^2)$ vertices (*i.e.* $|\mathcal{K}(H)| \in O(k^2)$), and then Lemma 1 enables us to conclude.

We say that a critical clique is affected if it contains an affected vertex and denote by A the set of the affected critical cliques. As each edge of F affects two vertices, we have that $|A| \leqslant 2k$. Since H is a 3-leaf power, its critical clique graph $\mathcal{C}(H)$ is a tree. Let T be the minimal subtree of $\mathcal{C}(H)$ that spans the affected critical cliques. Let us observe that if B is a maximal subtree of $\mathcal{C}(H) - T$, then none of the critical cliques in B contains an affected vertex and thus B was the critical clique graph of a 1-branch of G, which has been reduced by Rule 3 or Rule 4. Let $A' \subset \mathcal{K}(H)$ be the critical cliques of degree at least 3 in T. As

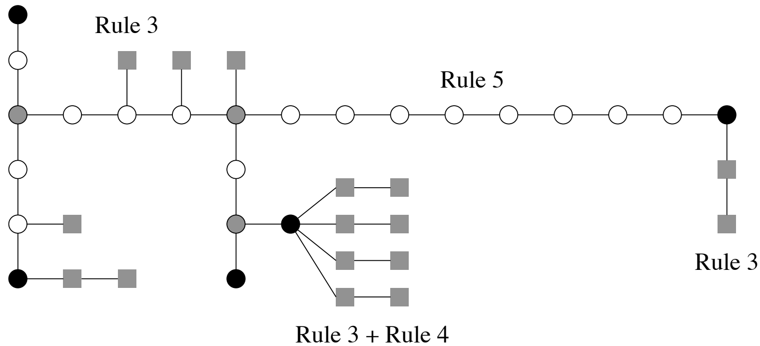

Fig. 3. The black circles are the critical cliques of A, the grey ones belong to A', and the squares are the critical cliques not in T. On the figure, we can observe a 2-branch of size 8 reduced by Rule 5. Application of Rule 3 may let a path of two critical cliques pendant to the elements of A and a single critical clique pendant to the elements of the small 2-branches. Finally, Rule 4 can only affect critical cliques of A.

$|A| \leqslant 2k$, we also have $|A'| \leqslant 2k$. The connected components resulting from the removal of A and A' in T are paths. There are at most $4k$ such paths. Each of these paths is composed by non-affected critical cliques. It follows that each of them corresponds to $path(B)$ for some 2-branch B of G, which has been reduced by Rule 5.

From these observations, we can now estimate the size of the reduced graph. Attached to each of the critical cliques of $T \setminus A$, we can have 1 pendant critical clique resulting from the application of Rule 3. Remark that any 2-branch reduced by Rule 5 has no such pendant clique and that $path(B)$ contains 5 critical cliques. So, a considered 2-branch in $\mathcal{C}(H)$ is made of at most 8 critical cliques. Finally, attached to each critical clique of A, we can have at most $(4k+2)$ extra critical cliques resulting from the application of Rule 4. See Figure 3 for an illustration of the shape of $\mathcal{C}(H)$. Summing up everything, we obtain that $\mathcal{C}(H)$ contains at most $4k \cdot 8 + 2k \cdot 2 + 2k \cdot (4k+3) = 8k^2 + 42k$ critical cliques.

By Lemma 1 we know that for each edited edge in a graph the number of critical cliques increases by at most 4. It follows that $\mathcal{K}(G)$ contains at most $8k^2 + 46k$ critical cliques, each of size at most $k+1$ (Rule 2). Thus, the reduced graph contains at most $8k^3 + 54k^2 + 46k$ vertices, proving the $O(k^3)$ kernel size. □

4 Kernels for Edge Completion and Edge Deletion

We now explain and adapt the previous rules to the cases where only insertions or only deletions of edges are allowed. First, observe that Rules 1 and 2 are also safe for 3-Leaf Power Completion and 3-Leaf Power Edge-Deletion (Rule 2 directly follows from Lemma 2). The same holds for Rules 3 and 4. However, this is not the case for the 2-branch reduction rule (Rule 5), which is safe for 3-Leaf Power Edge-Deletion, but not for 3-Leaf Power Completion.

Nevertheless, in the latter case, the following lemma yields a rule specific to the 3-LEAF POWER COMPLETION.

Lemma 6. *Let G be a graph admitting a clean 2-branch B such that path(B) is composed by at least $k+4$ critical cliques. If P_1 and P_2 belong to the same connected component in G, then there is no 3-leaf power completion of size at most k.*

Rule 6. *Let G be a graph having a clean 2-branch B with attachment points P_1 and P_2 such that path(B) is composed by at least $k+4$ critical cliques.*

- *If P_1 and P_2 belong to the same connected component in $G \setminus B^R$, then there is no completion of size at most k.*
- *Otherwise, remove from G all the vertices of $V(B)$ except those of $P_1 \cup Q_1 \cup P_2 \cup Q_2$ and add all possible edges between Q_1 and Q_2.*

Using Rules 3, 4 and 5 for deletion and Rules 3, 4 and 6 for completion, we obtain a kernel with $O(k^3)$ vertices.

Theorem 5. *The* PARAMETERIZED 3-LEAF POWER COMPLETION *and* PARAMETERIZED 3-LEAF POWER EDGE-DELETION *problems admit kernels with $O(k^3)$ vertices. Given a graph G, a reduced instance can be computed in linear time.*

References

1. Barthélémy, J.-P., Guénoche, A.: Trees and proximity representations. John Wiley & Sons, Chichester (1991)
2. Böcker, S., Briesemeister, S., Klau, G.W.: Exact algorithms for cluster editing: evaluation and experiments. In: McGeoch, C.C. (ed.) WEA 2008. LNCS, vol. 5038, pp. 289–302. Springer, Heidelberg (2008)
3. Brandstädt, A., Le., V.B.: Structure and linear time recognition of 4-leaf powers. ACM Transactions on Algorithms (to appear)
4. Brandstädt, A., Le, V.B.: Structure and linear time recognition of 3-leaf powers. Information Processing Letters 98(4), 133–138 (2006)
5. Chang, M.-S., Ko, M.-T.: The 3-steiner root problem. In: Brandstädt, A., Kratsch, D., Müller, H. (eds.) WG 2007. LNCS, vol. 4769, pp. 109–120. Springer, Heidelberg (2007)
6. Dehne, F., Langston, M., Luo, X., Pitre, S., Shaw, P., Zhang, Y.: The cluster editing problem: implementations and experiments. In: Bodlaender, H.L., Langston, M.A. (eds.) IWPEC 2006. LNCS, vol. 4169, pp. 13–24. Springer, Heidelberg (2006)
7. Dom, M., Guo, J., Hüffner, F., Neidermeier, R.: Error compensation in leaf root problems. Algorithmica 44, 363–381 (2006)
8. Dom, M., Guo, J., Hüffner, F., Niedermeier, R.: Error compensation in leaf root problems. In: Fleischer, R., Trippen, G. (eds.) ISAAC 2004. LNCS, vol. 3341, pp. 389–401. Springer, Heidelberg (2004)
9. Dom, M., Guo, J., Hüffner, F., Niedermeier, R.: Extending the tractability border for closest leaf powers. In: Kratsch, D. (ed.) WG 2005. LNCS, vol. 3787, pp. 397–408. Springer, Heidelberg (2005)

10. Dom, M., Guo, J., Hüffner, F., Niedermeier, R.: Closest 4-leaf power is fixed-parameter tractable. Discrete Applied Mathematics 156(18), 3345–3361 (2008)
11. Downey, R.G., Fellows, M.R.: Parameterized complexity. Springer, Heidelberg (1999)
12. Fellows, M.R., Langston, M., Rosamond, F., Shaw, P.: Efficient parameterized preprocessing for cluster editing. In: Csuhaj-Varjú, E., Ésik, Z. (eds.) FCT 2007. LNCS, vol. 4639, pp. 312–321. Springer, Heidelberg (2007)
13. Goldberg, P.W., Golumbic, M.C., Kaplan, H., Shamir, R.: Four strikes against physical mapping of DNA. Journal of Computational Biology 2, 139–152 (1995)
14. Gramm, J., Guo, J., Hüffner, F., Niedermeier, R.: Graph-modeled data clustering: Exact algorithms for clique generation. Theory of Computing Systems 38(4), 373–392 (2005)
15. Guo, J.: A more effective linear kernelization for cluster editing. In: Chen, B., Paterson, M., Zhang, G. (eds.) ESCAPE 2007. LNCS, vol. 4614, pp. 36–47. Springer, Heidelberg (2007)
16. Guo, J., Niedermeier, R.: Invitation to data reduction and problem kernelization. SIGACT News 38(1), 31–45 (2007)
17. Hansen, P., Jaumard, B.: Cluster analysis and mathematical programming. Math. Program. 79(1-3), 191–215 (1997)
18. Kaplan, H., Shamir, R., Tarjan, R.E.: Tractability of parameterized completion problems on chordal, strongly chordal, and proper interval graphs. SIAM J. Comput 28, 780–791 (1999)
19. Kearney, P., Corneil, D.: Tree powers. Journal of Algorithms 29(1), 111–131 (1998)
20. Lin, G.H., Kearney, P.E., Jiang, T.: Phylogenetic k-root and steiner k-root. In: Lee, D.T., Teng, S.-H. (eds.) ISAAC 2000. LNCS, vol. 1969, pp. 539–551. Springer, Heidelberg (2000)
21. Natazon, A., Shamir, R., Sharan, R.: Complexity classification of some edge modification problems. Discrete Applied Mathematics 113, 109–128 (2001)
22. Neidermeier, R.: Invitation to fixed parameter algorithms. Oxford Lectures Series in Mathematics and its Applications, vol. 31. Oxford University Press, Oxford (2006)
23. Nishimura, N., Ragde, P., Thilikos, D.: On graph powers for leaf-labeled trees. Journal of Algorithms 42(1), 69–108 (2002)
24. Protti, F., Dantas da Silva, M., Szwarcfiter, J.L.: Applying modular decomposition to parameterized cluster editing problems. Theory of Computing Systems (2007)
25. Shamir, R., Sharan, R., Tsur, D.: Cluster graph modification problems. Discrete Applied Mathematics 144(1-2), 173–182 (2004)
26. Tedder, M., Corneil, D., Habib, M., Paul, C.: Simpler linear-time modular decomposition via recursive factorizing permutations. In: Aceto, L., Damgård, I., Goldberg, L.A., Halldórsson, M.M., Ingólfsdóttir, A., Walukiewicz, I. (eds.) ICALP 2008, Part I. LNCS, vol. 5125, pp. 634–645. Springer, Heidelberg (2008)